

日本国特許庁  
JAPAN PATENT OFFICE

260  
村上  
Jc971 U.S. PTO  
10/058264  
01/29/02

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日  
Date of Application:

2001年 2月 1日

出願番号  
Application Number:

特願2001-025246

出願人  
Applicant(s):

日本電気株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年11月16日

特許庁長官  
Commissioner,  
Japan Patent Office

及川耕造



出証番号 出証特2001-3099848

【書類名】 特許願

【整理番号】 66206370

【提出日】 平成13年 2月 1日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 15/16

【発明者】

    【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

    【氏名】 富岡 耕治

【特許出願人】

    【識別番号】 000004237

    【氏名又は名称】 日本電気株式会社

【代理人】

    【識別番号】 100088959

    【弁理士】

    【氏名又は名称】 境 廣巳

【手数料の表示】

    【予納台帳番号】 009715

    【納付金額】 21,000円

【提出物件の目録】

    【物件名】 明細書 1

    【物件名】 図面 1

    【物件名】 要約書 1

    【包括委任状番号】 9002136

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 計算機システム及びCPU・メモリ搭載装置並びに入出力制御装置

【特許請求の範囲】

【請求項1】 それぞれ少なくとも1個のCPU及びメモリを有する複数のCPU・メモリ搭載装置と複数の入出力制御装置とをネットワークで相互に接続したクラスタ型の計算機システム。

【請求項2】 それぞれ少なくとも1個のCPU及びメモリを有する複数のCPU・メモリ搭載装置と、複数の入出力制御装置と、前記複数のCPU・メモリ搭載装置と前記複数の入出力制御装置とを相互に接続するネットワークとを備え、且つ、前記それぞれのCPU・メモリ搭載装置に、自装置の前記CPUから発行された入出力命令を自装置に事前に割り当てられた前記入出力制御装置に対して前記ネットワーク経由で送信すると共に前記入出力制御装置からの応答を前記ネットワーク経由で受信する通信手段を備え、前記それぞれの入出力制御装置に、自装置に事前に割り当てられた前記CPU・メモリ搭載装置からの入出力命令を前記ネットワーク経由で受信すると共に当該入出力命令の応答を前記CPU・メモリ搭載装置に対して前記ネットワーク経由で送信する通信手段を備えた計算機システム。

【請求項3】 前記入出力制御装置の前記通信手段は、前記ネットワーク経由で受信した入出力命令の送付元が事前に設定されたCPU・メモリ搭載装置である場合に限り当該入出力命令を有効なものとして受信する手段を備える請求項2記載の計算機システム。

【請求項4】 前記CPU・メモリ搭載装置の前記通信手段は、前記ネットワーク経由で受信した応答の送付元が事前に設定された入出力制御装置である場合に限り当該応答を有効なものとして受信する手段を備える請求項2または3記載の計算機システム。

【請求項5】 前記ネットワークを前記複数のCPU・メモリ搭載装置間の通信にも使用する構成を有する請求項2記載の計算機システム。

【請求項6】 前記CPU・メモリ搭載装置の前記通信手段は、前記ネット

ワーク経由で他のCPU・メモリ搭載装置と通信する手段を備えた請求項5記載の計算機システム。

【請求項7】 前記複数のCPU・メモリ搭載装置間の通信は、他のCPU・メモリ搭載装置に搭載されたメモリのメモリアクセスにかかる通信である請求項6記載の計算機システム。

【請求項8】 前記複数のCPU・メモリ搭載装置の何れかが障害により動作不能になったとき、動作不能になった前記CPU・メモリ搭載装置が使用していた前記入出力制御装置を他の正常な前記CPU・メモリ搭載装置に割り当ててシステムの運用を継続させる制御を行う手段を備えた請求項2乃至7の何れか1項に記載の計算機システム。

【請求項9】 前記他の正常なCPU・メモリ搭載装置として、他の入出力制御装置を使って処理を行っている現用系のCPU・メモリ搭載装置を使用する請求項8記載の計算機システム。

【請求項10】 予備のCPU・メモリ搭載装置を少なくとも1台備え、前記他の正常なCPU・メモリ搭載装置として、前記予備のCPU・メモリ搭載装置を使用する請求項8記載の計算機システム。

【請求項11】 予備の入出力制御装置を少なくとも1台備え、且つ、現用の前記複数の入出力制御装置の何れかが障害により動作不能になったとき、動作不能になった前記入出力制御装置を使用していた前記CPU・メモリ搭載装置に前記予備の入出力制御装置を割り当ててシステムの運用を継続させる制御を行う手段を備えた請求項2乃至7の何れか1項に記載の計算機システム。

【請求項12】 少なくとも1個のCPU及びメモリを有するCPU・メモリ搭載装置と、入出力制御装置と、前記CPU・メモリ搭載装置と前記入出力制御装置とを接続する通信ケーブルとを備え、且つ、前記CPU・メモリ搭載装置に、前記CPUから発行された入出力命令を前記入出力制御装置に対して前記通信ケーブル経由で送信すると共に前記入出力制御装置からの応答を前記通信ケーブル経由で受信する通信手段を備え、前記入出力制御装置に、前記CPU・メモリ搭載装置からの入出力命令を前記通信ケーブル経由で受信すると共に当該入出力命令の応答を前記CPU・メモリ搭載装置に対して前記通信ケーブル経由で送

信する通信手段を備えた計算機システム。

【請求項 1 3】 少なくとも 1 個の CPU 及びメモリと、外部との通信手段であって前記 CPU から発行された入出力命令を事前に割り当てられた入出力制御装置宛に送信すると共に前記入出力制御装置からの応答を受信する通信手段とを 1 つのボードに実装した CPU・メモリ搭載装置。

【請求項 1 4】 前記通信手段は、前記受信した応答の送付元が事前に設定された入出力制御装置である場合に限って当該応答を有効なものとして受信する手段を備える請求項 1 3 記載の CPU・メモリ搭載装置。

【請求項 1 5】 入出力命令に基づいて周辺機器を制御する入出力制御回路と、外部との通信手段であって事前に割り当てられた CPU・メモリ搭載装置からの入出力命令を受信して前記入出力制御回路に伝達すると共に当該入出力命令の応答を前記 CPU・メモリ搭載装置宛に送信する通信手段とを 1 つのボードに実装した入出力制御装置。

【請求項 1 6】 前記通信手段は、前記受信した入出力命令の送付元が事前に設定された CPU・メモリ搭載装置である場合に限って当該入出力命令を有効なものとして受信する手段を備える請求項 1 5 記載の入出力制御装置。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、計算機システム及び計算機システムを構成する CPU・メモリ搭載装置並びに入出力制御装置に関する。

【0 0 0 2】

【従来の技術】

従来の計算機システムの一例を図 1 3 に示す。同図に示す計算機システム 1 0 0 1 では、CPU バス 1 0 0 2 によって相互に接続された 4 台の CPU 1 0 0 3 及び制御回路 1 0 0 4 と、この制御回路 1 0 0 4 にメモリ信号線 1 0 1 1 を通じて接続されたメモリ 1 0 0 5、及び入出力信号線 1 0 1 2 を通じて接続された 2 台の入出力制御回路 1 0 0 6 とが 1 つのボード 1 0 1 3 上に搭載されており、入出力制御回路 1 0 0 6 に、P C I 等の標準入出力バス 1 0 0 7 及び周辺機器制御

機能を有する入出力カード1008を介して周辺機器1009が接続されている。また、診断制御回路（SP；サービスプロセッサ）1010があり、制御回路1004及び入出力制御回路1006の故障診断や各部の初期設定等がこの診断制御回路1010によって実行されるようになっている。

#### 【0003】

メモリ1005には、オペレーティングシステム（OS）やその他の各種のアプリケーションプログラムが記憶されており、4台のCPU1003は1つのOSの制御の下でアプリケーションプログラムを実行することにより、所定の業務処理を遂行する。メモリ1005には、また、処理対象となる各種のデータが記憶される。制御回路1004は、CPU1003とメモリ1005及び入出力制御回路1006との間にあって、CPU1003から出される入出力命令やメモリアクセス命令を入出力制御回路1006やメモリ1005に伝達する機能や、それらの応答をCPU1003に伝達する機能などを有する。入出力制御回路1006は、制御回路1004を通じてCPU1003から送られてきた入出力命令に基づいて、その配下の標準入出力バス1007及び入出力カード1008を通じて周辺機器1009を制御し、その入出力命令の処理結果である応答データを制御回路1004を通じてCPU1003に返却する機能などを持つ。

#### 【0004】

他方、最近の計算機システムは、主にシステム全体の可用性を向上させるために、複数の計算機システムをネットワークで接続したクラスタ構成が一般的になってきている。このようなクラスタ構成を採用した従来の計算機システムの構成例を図14に示す。図14に示したクラスタ計算機は、図13で説明した従来の計算機システム1001を3台、ネットワーク1020によって相互に接続したものである。この例では、それぞれの計算機システム1001におけるネットワーク接続用の入出力カード1021を使ってネットワーク接続しているが、それぞれの計算機システム1001の制御回路1004の部分でネットワーク接続する構成もある。

#### 【0005】

このようなクラスタ計算機では、それぞれの計算機システム1001は自シス

テム専用のOSの制御の下に1つの計算機として独立して動作し、また、ネットワーク1020を通じて計算機システム1001間で通信を行うことができる。このため、異なるジョブを別々の計算機システム1001で実行させたり、1つのジョブを並列プログラムとして同時に複数の計算機システム1001で実行させるなど、多様なジョブ処理形態を取ることができる。また、何れかの計算機システム1001がダウンしても、残りの正常な計算機システム1001で運用を継続することができる利点がある。

## 【0006】

## 【発明が解決しようとする課題】

従来の計算機システムが抱える問題点は、計算機システム1001を構成するCPU1003やメモリ1005に障害が発生して使用不可能になると、その計算機システム1001内の入出力制御回路1006に障害がなく正常であっても、その正常な入出力制御回路1006及びその配下の周辺機器1009も利用できなくなることである。その理由は、従来の計算機システム1001では、入出力制御回路1006はそれに制御回路1004を介して直結されたCPU1003からしか制御することができないためである。また、CPU1003とそれが発行する入出力命令を遂行する入出力制御回路1006とが同じボード1013上に組み付けられており、このボード1013が保守交換の最小単位になるからである。

## 【0007】

そこで本発明の目的は、計算機システムにおける障害時の可用性をより一層高めることにある。

## 【0008】

他方、1台の装置に直結されていた周辺装置などを共用化するためにネットワーク接続する考えは、特開2000-141831号公報、特開2000-172463号公報、特開2000-293341号公報など多数の文献に記載されている。しかしながら、CPUとそれが専用使用する入出力制御回路とをネットワーク接続した文献は見当たらない。その理由は、一般にネットワーク接続は共用化が目的であるため、CPUとそれが専用使用する入出力制御回路とをネ

ットワーク接続しても意味がないと考えられていたことによる。また、入出力制御回路をネットワーク接続すると、オペレーティングシステムの改造が必要になると考えられていたことも、理由の一つである。

## 【 0 0 0 9 】

そこで本発明の別の目的は、オペレーティングシステムの改造無しにCPUとそれが専用に使用する入出力制御装置とをネットワーク接続し得るようにすることにある。

## 【 0 0 1 0 】

更に、ネットワーク接続したときの問題として、ネットワーク上のどこからも基本的にアクセス可能であるため予期せぬ相手からの誤ったアクセスによって誤動作する危険性がある。

## 【 0 0 1 1 】

そこで本発明の更に別の目的は、事前に設定した相手以外からのアクセスをオペレーティングシステムの改造無しに禁止することによって、ネットワーク接続に伴う危険性を除去することにある。

## 【 0 0 1 2 】

本発明のその他の目的は、以下に述べる本発明の実施の形態の説明によって明らかになるであろう。

## 【 0 0 1 3 】

## 【課題を解決するための手段】

本発明を適用したクラスタ型の計算機システムは、それぞれ少なくとも1個のCPU及びメモリを有する複数のCPU・メモリ搭載装置と複数の入出力制御装置とがネットワークで相互に接続されている。より具体的には、本発明の第1の計算機システムは、それぞれ少なくとも1個のCPU及びメモリを有する複数のCPU・メモリ搭載装置と、複数の入出力制御装置と、前記複数のCPU・メモリ搭載装置と前記複数の入出力制御装置とを相互に接続するネットワークとを備え、且つ、前記それぞれのCPU・メモリ搭載装置に、自装置の前記CPUから発行された入出力命令を自装置に事前に割り当てられた前記入出力制御装置に対して前記ネットワーク経由で送信すると共に前記入出力制御装置からの応答を前



記ネットワーク経由で受信する通信手段を備え、前記それぞれの入出力制御装置に、自装置に事前に割り当てられた前記CPU・メモリ搭載装置からの入出力命令を前記ネットワーク経由で受信すると共に当該入出力命令の応答を前記CPU・メモリ搭載装置に対して前記ネットワーク経由で送信する通信手段を備えている。

## 【0014】

また、本発明の第2の計算機システムは、第1の計算機システムにおいて、前記入出力制御装置の前記通信手段は、前記ネットワーク経由で受信した入出力命令の送付元が事前に設定されたCPU・メモリ搭載装置である場合に限り当該入出力命令を有効なものとして受信する手段を備える。

## 【0015】

また、本発明の第3の計算機システムは、第1または第2の計算機システムにおいて、前記CPU・メモリ搭載装置の前記通信手段は、前記ネットワーク経由で受信した応答の送付元が事前に設定された入出力制御装置である場合に限り当該応答を有効なものとして受信する手段を備える。

## 【0016】

また、本発明の第4の計算機システムは、第1の計算機システムにおいて、前記ネットワークを前記複数のCPU・メモリ搭載装置間の通信にも使用する構成を有する。

## 【0017】

また、本発明の第5の計算機システムは、第4の計算機システムにおいて、前記CPU・メモリ搭載装置の前記通信手段は、前記ネットワーク経由で他のCPU・メモリ搭載装置と通信する手段を備える。

## 【0018】

また、本発明の第6の計算機システムは、第5の計算機システムにおいて、前記複数のCPU・メモリ搭載装置間の通信は、他のCPU・メモリ搭載装置に搭載されたメモリのメモリアクセスにかかる通信としている。

## 【0019】

また、本発明の第7の計算機システムは、第1乃至第6の計算機システムにお

いて、前記複数のCPU・メモリ搭載装置の何れかが障害により動作不能になったとき、動作不能になった前記CPU・メモリ搭載装置が使用していた前記入出力制御装置を他の正常な前記CPU・メモリ搭載装置に割り当ててシステムの運用を継続させる制御を行う手段を備える。

## 【0020】

また、本発明の第8の計算機システムは、第7の計算機システムにおいて、前記他の正常なCPU・メモリ搭載装置として、他の入出力制御装置を使って処理を行っている現用系のCPU・メモリ搭載装置を使用する。

## 【0021】

また、本発明の第9の計算機システムは、第7の計算機システムにおいて、予備のCPU・メモリ搭載装置を少なくとも1台備え、前記他の正常なCPU・メモリ搭載装置として、前記予備のCPU・メモリ搭載装置を使用する。

## 【0022】

また、本発明の第10の計算機システムは、第1乃至第6の計算機システムにおいて、予備の入出力制御装置を少なくとも1台備え、且つ、現用の前記複数の入出力制御装置の何れかが障害により動作不能になったとき、動作不能になった前記入出力制御装置を使用していた前記CPU・メモリ搭載装置に前記予備の入出力制御装置を割り当ててシステムの運用を継続させる制御を行う手段を備える。

## 【0023】

また、本発明の第11の計算機システムは、少なくとも1個のCPU及びメモリを有するCPU・メモリ搭載装置と、入出力制御装置と、前記CPU・メモリ搭載装置と前記入出力制御装置とを接続する通信ケーブルとを備え、且つ、前記CPU・メモリ搭載装置に、前記CPUから発行された入出力命令を前記入出力制御装置に対して前記通信ケーブル経由で送信すると共に前記入出力制御装置からの応答を前記通信ケーブル経由で受信する通信手段を備え、前記入出力制御装置に、前記CPU・メモリ搭載装置からの入出力命令を前記通信ケーブル経由で受信すると共に当該入出力命令の応答を前記CPU・メモリ搭載装置に対して前記通信ケーブル経由で送信する通信手段を備える。

## 【 0 0 2 4 】

他方、本発明の第 1 の CPU・メモリ搭載装置は、少なくとも 1 個の CPU 及びメモリと、外部との通信手段であって前記 CPU から発行された入出力命令を事前に割り当てられた入出力制御装置宛に送信すると共に前記入出力制御装置からの応答を受信する通信手段とを 1 つのボードに実装した構成を有する。

## 【 0 0 2 5 】

また、本発明の第 2 の CPU・メモリ搭載装置は、第 1 の CPU・メモリ搭載装置において、前記通信手段は、前記受信した応答の送付元が事前に設定された入出力制御装置である場合に限り当該応答を有効なものとして受信する手段を備える。

## 【 0 0 2 6 】

更に、本発明の第 1 の入出力制御装置は、入出力命令に基づいて周辺機器を制御する入出力制御回路と、外部との通信手段であって事前に割り当てられた CPU・メモリ搭載装置からの入出力命令を受信して前記入出力制御回路に伝達すると共に当該入出力命令の応答を前記 CPU・メモリ搭載装置宛に送信する通信手段とを 1 つのボードに実装した構成を有する。

## 【 0 0 2 7 】

また、本発明の第 2 の入出力制御装置は、第 1 の入出力制御装置において、前記通信手段は、前記受信した入出力命令の送付元が事前に設定された CPU・メモリ搭載装置である場合に限り当該入出力命令を有効なものとして受信する手段を備える。

## 【 0 0 2 8 】

## 【作用】

本発明を適用したクラスタ型の計算機システムにあっては、システムの立ち上げ時、CPU・メモリ搭載装置の通信手段に当該 CPU・メモリ搭載装置で使用する入出力制御装置の宛先情報を設定し、他方、その入出力制御装置の通信手段には当該入出力制御装置を使用する CPU・メモリ搭載装置の宛先情報を設定しておく。システムの運用が開始され、CPU・メモリ搭載装置上の CPU から入出力命令が発行されると、それが通信手段に渡され、通信手段がシステム立ち上

げ時に設定された宛先情報に従ってその入出力命令をネットワーク経由で該当する入出力制御装置へ送信し、入出力制御装置ではこの入出力命令を通信手段で受信してそれに基づいて周辺機器を制御する。そして、当該入出力命令の応答を入出力制御装置の通信手段がシステム立ち上げ時に設定された宛先情報に従ってネットワーク経由で該当するCPU・メモリ搭載装置に送信し、CPU・メモリ搭載装置ではこの応答を通信手段で受信し、CPUに伝達する。

## 【 0 0 2 9 】

このようにCPU・メモリ搭載装置においては、通信手段がCPUから発行された入出力命令の所定の宛先への送信とその応答の受信を担っているため、CPUは入出力命令の発行やその応答の受け取りに関して従来と何ら変わるところがなく、入出力制御装置があたかも直結されているかのように見える。従って、オペレーティングシステムに手を一切加える必要がない。そして、或る入出力制御装置を現に使用していたCPU・メモリ搭載装置に障害が起きて使用不能になった場合、診断制御回路等によって、当該入出力制御装置の通信手段における宛先情報を他の正常なCPU・メモリ搭載装置に変更すると共に、その正常なCPU・メモリ搭載装置の通信手段に当該入出力制御装置の宛先情報を設定することにより、その正常なCPU・メモリ搭載装置から当該入出力制御装置を利用することが可能となり、システムの可用性をより一層向上させることができる。また、予備の入出力制御装置を用意しておけば、何れかの入出力制御装置に障害が発生して使用不能になったとき、それを使用していたCPU・メモリ搭載装置に前記予備の入出力制御装置を割り当てることで、当該CPU・メモリ搭載装置における入出力処理を伴う業務の継続が可能になる。

## 【 0 0 3 0 】

## 【発明の実施の形態】

次に本発明の実施の形態の例について図面を参照して詳細に説明する。

## 【 0 0 3 1 】

図1は本発明を適用した計算機システムで使用されるCPU・メモリ搭載装置の一例を示すブロック図である。この例のCPU・メモリ搭載装置101は、CPUバス102によって相互に接続された4台のCPU103及び制御回路10

4と、この制御回路104にメモリ信号線105を通じて接続されたメモリ106と、制御回路104に2組の入出力信号線107、108を通じて接続された通信回路109と、この通信回路109に接続された通信ケーブル接続用の接続具（例えばコネクタ）110とが、1つのボード上に実装されている。使用時には、接続具110に通信ケーブル111が接続される。入出力信号線が107と108の2組設けられているのは、このCPU・メモリ搭載装置101が図13の従来の計算機システム1001と同じく2つの入出力制御回路を制御できるようになっているためである。入出力信号線107が入出力ポート0、入出力信号線108が入出力ポート1に対応する。

#### 【0032】

メモリ106は、例えばROM及びRAMで構成され、オペレーティングシステム（OS）やその他各種のアプリケーションプログラム及び処理対象となる各種データを記憶する。4台のCPU103は、メモリ106に記憶されたOSの制御の下にアプリケーションプログラムを実行することにより、所定の業務処理を遂行する。入出力命令を発行する際の動作及びメモリアクセスを行う際の動作は従来の計算機システムと全く同じであり、入出力命令及びメモリアクセス命令をCPUバス102上に出力する。入出力命令の場合、その入出力ポートを指定する。

#### 【0033】

制御回路104は、CPU103とメモリ106及び通信回路109との間の命令やデータの授受を中継する制御を行う。本例の場合、制御回路104は、CPU103からCPUバス102に入出力命令が発行されると、それを取り込んで2本ある入出力信号線107、108のうち、指定された入出力ポートの入出力信号線を通じて通信回路109へ伝達する。また、通信回路109から入出力信号線107、108を通じて入出力命令に対する応答を受信すると、それをCPUバス102を通じてCPU103に伝達する。他方、CPU103からCPUバス102にメモリアクセス命令が発行されると、それを取り込んでメモリ信号線105を通じてメモリ106に渡し、メモリに対するリード、ライトの動作を行わせる。リードデータ等のメモリアクセス命令に対する応答がメモリ106

からメモリ信号線105を通じて送られてくると、それをCPUバス102を通じてCPU103に返却する。

#### 【0034】

入出力信号線107、108の先には、図13に示した従来の計算機システム1001では入出力制御回路1006が接続されていたが、本例ではそれに代えて通信回路109が接続されている。この通信回路109には事前に送信先の宛先情報が設定されており、入出力信号線107、108から入出力命令を受信すると、この入出力命令に宛先情報を付加した通信メッセージを接続具110に接続された通信ケーブル111に出力する機能と、この出力した入出力命令に対する応答を含む通信メッセージを通信ケーブル111から受信すると、通信メッセージ中の応答を取り出して入出力信号線107、108に出力する機能とを備えている。

#### 【0035】

図2に通信回路109で行われる変換処理の概要を示す。一般にCPU103が発行する入出力命令、従って入出力信号線107、108を通じて制御回路104から通信回路109へ渡される入出力命令は、図2の符号121に示すように、入出力命令の種別等を示すI/Oコマンド122とそれに付随するデータ123とから構成される。データ123は一般的に出力データであり、従って入力命令のようにデータ123が付随しない場合もある。通信回路109は、このような入出力命令121を受けると、図2の符号131に示すように、I/Oコマンド122及びデータ123をそのままそっくり通信データ132として含み、それにネットワークコマンド133、送付先ID134及び送付元ID135を付加した通信メッセージに変換して通信ケーブル111に出力する。送付先ID134及び送付元ID135にはそれぞれ事前に設定されたIDが設定される。IDは、上位数ビットがノード番号を示し、残りの下位数ビットがノード内番号を示す。ネットワークコマンド133にはメッセージ長等の情報が設定されるが、具体的な中身は使用する通信方式に依存する。

#### 【0036】

他方、通信ケーブル111から受信する通信メッセージは、図2の符号141

に示すようにネットワークコマンド143、送付先ID144、送付元ID145及び通信データ142から構成され、形式的には通信メッセージ131と同じであるが、通信データ142の部分には、先に発行した入出力命令の応答が含まれている。通信回路109では、通信メッセージ141を受信すると、その中の通信データ142を抽出して図2の符号151に示すような形式の応答に変換して入出力信号線107、108に出力する。応答151は、通信データ142に含まれていた応答コマンド152及びそれに付随するデータ153から構成される。

#### 【0037】

以上のような機能を持つ通信回路109の構成例を図3に示す。この例の通信回路109は、図2で説明したような変換処理を行うポート0用の変換部161及びポート1用の変換部162と、それぞれの変換部161、162で作成された通信メッセージを順次に選択して通信ケーブル111に出力する選択部163と、通信ケーブル111から通信メッセージを受信し、その送付先ID144に基づいてポート0用の変換部161、ポート1用の変換部162に通信メッセージを振り分ける振分部164と、ポート0用の変換部161に振り分ける通信メッセージが持つべき送付先ID144の値及びポート1用の変換部162に振り分ける通信メッセージが持つべき送付先ID144の値を振分部164に提供する振分情報レジスタ165とを備えている。

#### 【0038】

また、ポート0用の変換部161は、通信相手を一意に識別するためのIDを保持する相手IDレジスタ171と、通信の送り主を一意に識別するためのIDを保持する自IDレジスタ172と、入出力信号線107から図2に示した入出力命令121を受信したとき、通信データ132に入出力命令121を、送付先ID134に相手IDレジスタ171の設定値を、送付元ID135に自IDレジスタ172の設定値をそれぞれ設定し、更にネットワークコマンド133を付加した通信メッセージ131を作成して、選択部163に出力するI/O→N変換部173と、その逆に、振分部164から図2に示した通信メッセージ141を受信したとき、それに含まれる通信データ142を抽出し、応答151として

入出力信号線107に出力するN→I/O変換部174とを備えている。

【0039】

相手IDレジスタ171には、当該CPU・メモリ搭載装置101が使用する入出力制御装置を一意に識別するノード番号とその入出力制御装置における入出力制御回路を一意に識別するノード内番号とがそれぞれ上位数ビット、下位数ビットに設定され、自IDレジスタ172には、当該CPU・メモリ搭載装置101を一意に識別するノード番号と当該CPU・メモリ搭載装置101における入出力ポート0を一意に識別するノード内番号とがそれぞれ上位数ビット、下位数ビットに設定される。

【0040】

また、本例のN→I/O変換部174は、通信メッセージ141中の送付元ID145と相手IDレジスタ171の設定値とを比較し、一致しなければ変換処理を実施せず、従って応答151を入出力信号線107に出力しない機能を有している。これは、図13で説明した従来の計算機システム1001と異なり、本例のCPU・メモリ搭載装置101は、後述するようにネットワークに接続された全ての入出力制御装置と基本的には通信可能であるが、実際のシステム運用においては、CPU・メモリ搭載装置101は事前に設定された入出力制御装置だけを使用する必要がある。従って、事前に設定された入出力制御装置以外の入出力制御装置から誤って応答が送られてきた際には、エラーとしてそれをCPUに伝達しない機構が必要である。N→I/O変換部174に設けられた上記のチェック機構はこのような役割を持つ。

【0041】

ポート1用の変換部162も、ポート0用の変換部161と同様な構成要素181～184で構成されている。

【0042】

図4は本発明を適用した計算機システムで使用される入出力制御装置の一例を示すブロック図である。この例の入出力制御装置201は、1台の入出力制御回路202と、この入出力制御回路202にPCI等の標準入出力バス203を介して接続された周辺機器制御機能を有する複数の入出力カード204と、入出力



制御回路 2 0 2 に入出力信号線 2 0 5 を通じて接続された通信回路 2 0 6 と、この通信回路 2 0 6 に接続された通信ケーブル接続用の接続具（例えばコネクタ）2 0 7 とが、1 つのボード上に実装されている。使用時には、接続具 2 0 7 に通信ケーブル 2 1 1 が接続される。また、入出力カード 2 0 4 にはそのカード種別に応じた周辺機器 2 1 2 が接続される。

#### 【 0 0 4 3 】

入出力制御回路 2 0 2 は、入出力信号線 2 0 5 から受信した入出力命令に基づいて、その配下の標準入出力バス 2 0 3 及び入出力カード 2 0 4 を通じて周辺機器 2 1 2 を制御し、その入出力命令の処理結果である応答を入出力信号線 2 0 5 に出力する機能を持ち、その構成と動作は図 1 3 に示した入出力制御回路 1 0 0 6 と何ら変わるところがない。但し、入出力信号線 2 0 5 の先には図 1 3 に示した従来の計算機システム 1 0 0 1 では制御回路 1 0 0 4 が接続されていたが、本例では、その部分に通信回路 2 0 6 が設けられている。

#### 【 0 0 4 4 】

通信回路 2 0 6 は、通信ケーブル 2 1 1 から通信メッセージを受信すると、通信メッセージ中の入出力命令を取り出して入出力信号線 2 0 5 に出力する機能と、入出力信号線 2 0 5 から入出力命令の応答を受信すると、その応答に事前に設定された宛先情報を付加した通信メッセージを通信ケーブル 2 1 1 に出力する機能とを備えている。この通信回路 2 0 6 で行われる変換処理の概要を先の図 2 を参照して説明する。

#### 【 0 0 4 5 】

通信ケーブル 2 1 1 から受信する通信メッセージは、図 2 の符号 1 3 1 に示した形式となる。通信回路 2 0 6 は、この通信メッセージ 1 3 1 を受信すると、通信データ 1 3 2 を抽出して図 2 の符号 1 2 1 に示した I / O コマンド 1 2 2 及びそれに付随するデータ 1 2 3 から構成される入出力命令 1 2 1 を生成し、それを入出力信号線 2 0 5 を通じて入出力制御回路 2 0 2 へ出力する。他方、入出力信号線 2 0 5 を通じて入出力制御回路 2 0 2 から出力される応答は、図 2 の符号 1 5 1 に示したような形式となる。通信回路 2 0 6 は、この応答 1 5 1 を受信すると、図 2 の符号 1 4 1 に示すように、応答コマンド 1 5 2 及びデータ 1 5 3 をそ

のままそっくり通信データ142として含み、それにネットワークコマンド143、送付先ID144及び送付元ID145を付加した通信メッセージ141に変換して通信ケーブル211に出力する。送付先ID144及び送付元ID145にはそれぞれ事前に設定されたIDが設定される。ネットワークコマンド143にはメッセージ長等の情報が設定されるが、具体的な中身は使用する通信方式に依存する。

## 【0046】

以上のような機能を持つ通信回路206は、図3に示したCPU・メモリ搭載装置側の通信回路109と同様に構成できる。この場合、本例の入出力制御装置201は入出力制御回路202を1台しか有していないので、図3の変換部161、162に相当する部分が1つだけ有れば足りる。その構成例を図5に示す。この例の通信回路206は、通信相手を一意に識別するためのIDを保持する相手IDレジスタ221と、通信の送り主を一意に識別するためのIDを保持する自IDレジスタ222と、通信ケーブル211から図2に示した通信メッセージ131を受信したとき、それに含まれる通信データ132を抽出し、入出力命令121として入出力信号線205に出力するN→I/O変換部223と、その逆に、入出力信号線205から図2に示した応答151を受信したとき、通信データ142に応答151を、送付先ID144に相手IDレジスタ221の設定値を、送付元ID145に自IDレジスタ222の設定値をそれぞれ設定し、更にネットワークコマンド143を付加した通信メッセージ141を作成して、通信ケーブル211に出力するI/O→N変換部224とを備えている。

## 【0047】

相手IDレジスタ221には、当該入出力制御装置201を使用するCPU・メモリ搭載装置を一意に識別するノード番号とそのCPU・メモリ搭載装置におけるノード内番号とがそれぞれ上位数ビット、下位数ビットに設定され、自IDレジスタ222には、当該入出力制御装置201を一意に識別するノード番号と当該入出力制御装置201における入出力制御回路202を一意に識別するノード内番号とがそれぞれ上位数ビット、下位数ビットに設定される。

## 【0048】

また、本例のN→I/O変換部223は、通信メッセージ131中の送付元ID135と相手IDレジスタ221の設定値とを比較し、一致しなければ変換処理を実施せず、従って入出力命令121を入出力信号線205に出力しない機能を有している。これは、図13で説明した従来の計算機システム1001と異なり、本例の入出力制御装置201は、後述するようにネットワークに接続された全てのCPU・メモリ搭載装置と基本的には通信可能であるが、実際のシステム運用においては、入出力制御装置201は事前に設定されたCPU・メモリ搭載装置だけで使用される必要がある。従って、事前に設定されたCPU・メモリ搭載装置以外のCPU・メモリ搭載装置から入出力命令を含む通信メッセージが誤って送られてきた際には、エラーとしてそれを入出力制御回路202に伝達しない機構が必要である。N→I/O変換部223に設けられた上記のチェック機構はこのような役割を持つ。

## 【0049】

図6は本発明を適用した計算機システムの一例を示すブロック図である。この例の計算機システムは、図1に示したCPU・メモリ搭載装置101及び図4に示した入出力制御装置201をそれぞれ3台使用して、クラスタ型の計算機システムを実現している。図6では、それぞれのCPU・メモリ搭載装置、入出力制御装置を区別するために枝番を用いて、101-1～101-3、201-1～201-3の符号を付けてある。

## 【0050】

各CPU・メモリ搭載装置101-1～101-3は、通信ケーブル111によってネットワーク301の接続点302～304に接続され、各入出力制御装置201-1～201-3は、通信ケーブル211によってネットワーク301の接続点305～307に接続されている。ネットワーク301には、各接続点302～307の宛先情報として、その接続点に接続されているCPU・メモリ搭載装置101-1～101-3及び入出力制御装置201-1～201-3のノード番号が事前に設定されており、各接続点302～307から流入する図2の通信メッセージ131、141をその送付先ID134、144中のノード番号と同じノード番号を持つ接続点に接続されたノード（CPU・メモリ搭載装置

、入出力制御装置)に届ける。このようなネットワーク301としては、例えばトーラス型ネットワーク、メッシュ型ネットワーク、クロスバ型ネットワーク等の高速ネットワークが使用される。

#### 【0051】

このように本例の計算機システムは、複数のCPU・メモリ搭載装置101-1~101-3及び複数の入出力制御装置201-1~201-3がネットワーク301で相互に接続されているため、基本的に、任意のCPU・メモリ搭載装置101-1~101-3に任意の入出力制御装置201-1~201-3を割り当てることができる。しかし、実際のシステム運用中は、各CPU・メモリ搭載装置101-1~101-3に特定の入出力制御装置201-1~201-3を論理的に割り当て、それらで情報処理装置を構成する。図6では、CPU・メモリ搭載装置101-1に入出力制御装置201-1を、CPU・メモリ搭載装置101-2に入出力制御装置201-2を、CPU・メモリ搭載装置101-3に入出力制御装置201-3をそれぞれ割り当てた例を示しており、CPU・メモリ搭載装置101-1と入出力制御装置201-1で1つの情報処理装置308が構成され、CPU・メモリ搭載装置101-2と入出力制御装置201-2で別の1つの情報処理装置309が構成され、CPU・メモリ搭載装置101-3と入出力制御装置201-3で更に別の1つの情報処理装置310が構成されている。

#### 【0052】

各CPU・メモリ搭載装置101-1~101-3への入出力制御装置201-1~201-3の割り当ては、それらの通信回路109、206における相手IDの設定によって行われる。具体的には、CPU・メモリ搭載装置101-1の通信回路109では、図3のポート0用の変換部161における相手IDレジスタ171に入出力制御装置201-1のノード番号及びノード内番号を設定し、入出力制御装置201-1の通信回路206では、図5の相手IDレジスタ221にCPU・メモリ搭載装置101-1のノード番号及びノード内番号を設定する。同様に、CPU・メモリ搭載装置101-2、101-3の通信回路109のポート0用の変換部161における相手IDレジスタ171に入出力制御装

置201-2、201-3のノード番号及びノード内番号を設定し、入出力制御装置201-2、201-3の通信回路206の相手IDレジスタ221にCPU・メモリ搭載装置101-2、101-3のノード番号及びノード内番号を設定する。CPU・メモリ搭載装置101-1~101-3の通信回路109におけるポート1用の変換部162はシステムの運用開始時点では使用しないため、相手IDは設定されない。

## 【0053】

上述のような通信回路109、206への相手IDの設定は、通信回路109、206への自IDの設定や振分情報の設定を含め、本例の計算機システムでは、各情報処理装置308~310毎に設けた診断制御回路(SP)311~313で行うようにしている。つまり、診断制御回路311~313に通信回路109、206へのパスを設け、このパスを通じて図3の相手IDレジスタ171、181、自IDレジスタ172、182及び振分情報レジスタ165の設定、及び図5の相手IDレジスタ221、自IDレジスタ222の設定を行う。また、各診断制御回路311~313は、図13で示した従来の計算機システム1001の診断制御回路1010と同様の機能を併せ持ち、また診断用ネットワーク314を通じて相互に通信可能になっている。

## 【0054】

更に、本例の計算機システムは、情報処理装置308~310相互間の通信を可能にするために、図14に示した従来の計算機システムと同様に、それぞれの情報処理装置308~310の入出力制御装置201-1~201-3におけるネットワーク接続用の入出力カード315を使ってネットワーク316で相互に接続してある。ネットワーク316としては、例えばイーサネット等が使用される。勿論、それぞれの情報処理装置308~310のCPU・メモリ搭載装置101-1~101-3における制御回路104の部分でネットワーク接続する構成も採用することができる。更に、ネットワーク301を利用することもできるが、これについては後述する。

## 【0055】

次に本例の計算機システムの動作を説明する。

## 【 0 0 5 6 】

図 6 の計算機システムのシステム立ち上げ時、診断制御回路 3 1 1 ~ 3 1 3 に  
よってシステム各部の初期設定が実施される。そして、その一環として前述した  
ように各通信回路 1 0 9、2 0 6 への相手 ID、自 ID、振分情報の設定が実施  
される。また、各 CPU・メモリ搭載装置 1 0 1 - 1 ~ 1 0 1 - 3 の入出力ポー  
ト 0 には 1 台の入出力制御装置 2 0 1 - 1 ~ 2 0 1 - 3 が接続されていること、  
入出力ポート 1 には入出力制御装置が接続されていないといったシステムの構成  
情報が初期設定される。従って、各 CPU・メモリ搭載装置 1 0 1 - 1 ~ 1 0 1  
- 3 の CPU 1 0 3 は入出力命令を発行する際、入出力ポート 0 を使用する。

## 【 0 0 5 7 】

何れかの CPU・メモリ搭載装置、例えば CPU・メモリ搭載装置 1 0 1 - 1  
の CPU 1 0 3 から入出力ポート 0 への入出力命令が発行されると、制御回路 1  
0 4 がそれを通信回路 1 0 9 へ伝達する。通信回路 1 0 9 は、図 3 のポート 0 用  
の変換部 1 6 1 における I/O → N 変換部 1 7 3 によって図 2 で説明したように  
入出力命令 1 2 1 を通信メッセージ 1 3 1 に変換して選択部 1 6 3、接続具 1 1  
0、通信ケーブル 1 1 1 を通じてネットワーク 3 0 1 の接続点 3 0 2 へ送出する  
。このときの通信メッセージ 1 3 1 の送付先 ID 1 3 4 は入出力制御装置 2 0 1  
- 1 を一意に識別するノード番号及び入出力制御回路 2 0 2 を一意に識別するノ  
ード内番号を示し、送付元 ID 1 3 5 は CPU・メモリ搭載装置 1 0 1 - 1 を一  
意に識別するノード番号及びポート 0 を一意に識別するノード内番号を示してい  
る。

## 【 0 0 5 8 】

ネットワーク 3 0 1 では、接続点 3 0 2 に流入した通信メッセージ 1 3 1 をそ  
の送付先 ID 1 3 4 中のノード番号に従って接続点 3 0 5 へ届け、入出力制御装  
置 2 0 1 - 1 に送り込む。入出力制御装置 2 0 1 - 1 の通信回路 2 0 6 は、この  
通信メッセージ 1 3 1 を通信ケーブル 2 1 1 及び接続具 2 0 7 を通じて受信し、  
図 5 の N → I/O 変換部 2 2 3 により送付元 ID 1 3 5 が相手 ID レジスタ 2 2  
1 に設定された相手 ID と一致することを確認して、図 2 で説明したように元の  
入出力命令 1 2 1 に変換し、入出力制御回路 2 0 2 に出力する。入出力制御回路

202は、その入出力命令121を受け取り解釈した上で、若し周辺機器212にかかる入出力命令であれば入出力カード204を通じて周辺機器212へその命令を伝達する等の制御を行い、周辺機器212は命令を解釈し、必要な入出力命令を実行した後に、結果を入出力制御回路202に返却する。他方、ネットワーク316経由による他の情報処理装置との間の通信にかかる入出力命令であった場合には、入出力制御回路202は入出力カード315へその命令を伝達する等の制御を行い、入出力カード315はその命令を解釈し、ネットワーク316経由で他の情報処理装置と通信を行い、結果を入出力制御回路202に返却する。

## 【0059】

入出力制御回路202は、返却された結果を入出力命令に対する応答として通信回路206に伝達する。通信回路206は、図5のI/O→N変換部224により、この応答を図2で説明したように応答151から通信メッセージ141に変換してネットワーク301の接続点305へ送り出す。このときの通信メッセージ141の送付先ID144は、CPU・メモリ搭載装置101-1を一意に識別するノード番号及びポート0を一意に識別するノード内番号であり、送付元ID145は入出力制御装置201-1を一意に識別するノード番号及び入出力制御回路202を一意に識別するノード内番号である。

## 【0060】

ネットワーク301は、接続点305に流入した通信メッセージ141をその送付先ID144中のノード番号に従って接続点302へ届け、CPU・メモリ搭載装置101-1に送り込む。CPU・メモリ搭載装置101-1の通信回路109は、この通信メッセージ141を図3の通信ケーブル111、接続具110を通じて振分部164で受信し、通信メッセージ141の送付先ID144と振分情報レジスタ165の振分情報とに基づいて、ポート0用の変換部161へ当該通信メッセージ141を振り分ける。変換部161は、N→I/O変換部174により、送付元ID145が相手IDレジスタ171に設定された相手IDと一致することを確認して、図2で説明したように元の応答151に変換し、制御回路104に出力する。制御回路104は、その応答151をCPU103に

通知する。

【0061】

他のCPU・メモリ搭載装置101-2、101-3上のCPU103から入出力命令が発行された際にも、前述と同様の動作を行われる。但し、CPU・メモリ搭載装置101-2から発行された入出力命令は、ネットワーク301経由で入出力制御装置201-2に伝えられて処理され、CPU・メモリ搭載装置101-3から発行された入出力命令は、ネットワーク301経由で入出力制御装置201-3に伝えられて処理されることになる。

【0062】

さて、システムの運用中に何れかのCPU・メモリ搭載装置、例えばCPU・メモリ搭載装置101-1に障害が発生し、動作の継続が不可能になったものとする、それが使用していた入出力制御装置201-1は自らは障害がないのに入出力命令を出すものがなくなるために、動作ができなくなる。診断制御回路311はこのような状態を検出すると、他の診断制御回路312、313と協調して、正常な入出力制御装置201-1の有効利用を図る。先ず、診断制御回路311は、入出力制御装置201-1を新たに使用するCPU・メモリ搭載装置を決定する。これは、どのCPU・メモリ搭載装置に障害が発生したとき、それが使用していた入出力制御装置を他のどのCPU・メモリ搭載装置が引き取るかの情報を事前に診断制御装置311～313に設定しておいて、その情報に基づいて決定しても良いし、他の診断制御回路とネゴシエーションを行って決定しても良い。以下では、CPU・メモリ搭載装置101-2が入出力制御装置201-1を使用するものと決定された場合を想定して、以後の動作を説明する。

【0063】

診断制御回路311は、入出力制御装置201-1の新たな使用元がCPU・メモリ搭載装置101-2に決定すると、入出力制御装置201-1の再立ち上げを実施し、その一環として、通信回路206の図5に示した相手IDレジスタ221にCPU・メモリ搭載装置101-2のノード番号及びポート1を示すノード内番号を設定し、自IDレジスタ222には元と同じIDを設定する。

【0064】



他方、CPU・メモリ搭載装置101-2側の診断制御回路312は、CPU・メモリ搭載装置312及び入出力制御装置201-2で仕掛かり中の処理が終了した時点で、情報処理装置309のシステム再立ち上げを実施し、その一環として、通信回路109の図3に示したポート1用の変換部162における相手IDレジスタ181に入出力制御装置201-1のノード番号及びノード内番号を、自IDレジスタ182にCPU・メモリ搭載装置101-2のノード番号及びポート1を示すノード内番号をそれぞれ設定する。また、振分情報レジスタ165に所望の振分情報を設定し、通信メッセージ中の送付先ID144のノード内番号がポート0を示す場合にはポート0の変換部161へ、ポート1を示す場合にはポート1の変換部162へ、それぞれ通信メッセージが振り分けられるようにする。ポート0用の変換部161の相手IDレジスタ171、自IDレジスタ172、入出力制御装置201-2の通信回路206における相手IDレジスタ221、自IDレジスタ222には元と同じIDを設定する。更に、CPU・メモリ搭載装置101-2の入出力ポート0には1台の入出力制御装置201-2が接続され、入出力ポート1には1台の入出力制御装置201-1が接続されているといったシステムの構成情報を設定する。従って、CPU・メモリ搭載装置101-2のCPU103は入出力命令を発行する際、入出力ポート0及び入出力ポート1の双方を使用することができる。

#### 【0065】

その後、システムの運用が再開され、CPU・メモリ搭載装置101-2のCPU103から入出力ポート1に入出力命令が発行されると、通信回路109のポート1用の変換部162により入出力制御装置201-1のノード番号を含む送付先IDを付加した通信メッセージに変換されて、ネットワーク301経由で入出力制御装置201-1に送られることになる。また、入出力制御装置201-1からの応答は、通信回路206によりCPU・メモリ搭載装置101-2のノード番号及びポート1を示すノード内番号を含む送付先IDを付加した通信メッセージに変換されて、ネットワーク301経由でCPU・メモリ搭載装置101-2に送られ、通信回路109のポート1用の変換部162により元の応答に変換されてCPU103に通知されることになる。

## 【 0 0 6 6 】

図 7 は本発明を適用した計算機システムの別のブロック図である。この例の計算機システムは、図 1 に示した CPU・メモリ搭載装置 1 0 1 と図 4 に示した入出力制御装置 2 0 1 をそれぞれ 1 台使用し、両者の接続具 1 1 0 と接続具 2 0 7 間を通信ケーブル 4 0 1 で直接接続することにより、CPU・メモリ搭載装置及び入出力制御装置をそれぞれ 1 台だけ有する最小構成の情報処理装置 4 0 2 を実現している。

## 【 0 0 6 7 】

このような最小構成の情報処理装置 4 0 2 は、性能や可用性の点ではクラスタ型に劣るが、価格が安いいため小規模な情報処理用に良く利用される。この場合、CPU・メモリ搭載装置 1 0 1 と入出力制御装置 2 0 1 とは 1 対 1 に対応するため、制御回路 1 0 4 と入出力制御回路 2 0 2 とを直結することも考えられるが、図 6 に示したクラスタ型の計算機システムで使う CPU・メモリ搭載装置及び入出力制御装置をそのままそっくり利用できるため、最小構成専用の設計が不要になり、部品点数も削減される利点がある。また、処理能力が不足してきた場合に、CPU・メモリ搭載装置及び入出力制御装置を買い足して、図 6 に示すようなクラスタ型に再構築できる利点もある。

## 【 0 0 6 8 】

図 7 の情報処理装置 4 0 2 における動作は、CPU・メモリ搭載装置 1 0 1 と入出力制御装置 2 0 1 との間でネットワーク経由でなく通信ケーブル 4 0 1 を介して通信メッセージが直接授受される点を除き、図 6 の計算機システムと同じである。

## 【 0 0 6 9 】

図 8 は本発明を適用した計算機システムの更に別のブロック図である。この例の計算機システムは、図 6 の計算機システムにおける CPU・メモリ搭載装置 1 0 1 - 1 ~ 1 0 1 - 3 間の通信をネットワーク 3 1 6 でなく、CPU・メモリ搭載装置 1 0 1 - 1 ~ 1 0 1 - 3 と入出力制御装置 2 0 1 - 1 ~ 2 0 1 - 3 間を相互に接続するネットワーク 3 0 1 を通じて行えるように変更したものである。そのために、CPU・メモリ搭載装置は図 1 の CPU・メモリ搭載装置 1 0 1 でな

く、図9に示されるCPU・メモリ搭載装置101Aを使用している。

【0070】

図9のCPU・メモリ搭載装置101Aが図1のCPU・メモリ搭載装置101と相違するところは、制御回路104A及び通信回路109Aの機能と、制御回路104Aと通信回路109Aとの間に新たにノード間通信バス112を設けた点にある。

【0071】

制御回路104Aは、CPU103からCPUバス102にメモリアクセス命令が出されたとき、そのアクセス先のメモリが自ノード上のメモリ106か、他ノード上のメモリかをアクセスアドレスに基づいて判断し、自ノード上のメモリ106のときは図1の制御回路104と同様にメモリ信号線105を通じて自ノードのメモリ106にメモリアクセス命令を伝達するが、他ノード上のメモリのときはノード間通信バス112を通じて通信回路109Aに伝達する。アクセスアドレスに基づいて、アクセス先が自ノード上のメモリ106か、どの他ノード上のメモリかを判断する方法としては、例えば、メモリアドレスの上位数ビットをノード番号に対応させ、当該上位数ビットが自ノードのノード番号と等しければ自ノードのメモリ106へのアクセス、それ以外は当該上位数ビットで示されるノード番号に対応するノード上のメモリへのアクセスと判断する方法などを採用することができる。また、制御回路104Aは、ノード間通信バス112を通じて通信回路109Aから前記メモリアクセス命令の応答を受信すると、それをCPU103へ伝達する。更に制御回路104Aは、ノード間通信バス112を通じて通信回路109Aから他ノードが発したメモリアクセス命令を受信すると、それをメモリ信号線105を通じてメモリ106に送ってアクセスを行わせ、その結果をノード間通信バス112を通じて通信回路109Aに伝達する。それ以外の機能は、図1の制御回路104と同じである。

【0072】

通信回路109Aは、図1の通信回路109の機能に加えて、ノード間通信バス112からメモリアクセス命令を受信すると、このメモリアクセス命令を含む該当ノード宛の通信メッセージを作成して通信ケーブル111に出力する機能、

このメモリアクセス命令に対する応答を含む通信メッセージを通信ケーブル 111 から受信すると、通信メッセージ中の応答を取り出してノード間通信パス 112 に出力する機能、通信ケーブル 111 から他ノードが出したメモリアクセス命令を含む通信メッセージを受信すると、通信メッセージ中のメモリアクセス命令を取り出してノード間通信パス 112 に出力する機能、そのメモリアクセス命令に対する応答をノード間通信パス 112 から受信すると、その応答を含む該当ノード宛の通信メッセージを作成して通信ケーブル 111 に出力する機能を有している。

### 【0073】

図 10 に通信回路 109A で行われるメモリアクセス関連の変換処理の概要を示す。一般に CPU 103 が発行するメモリアクセス命令、従ってノード間通信パス 112 を通じて制御回路 104A から通信回路 109A へ渡されるメモリアクセス命令は、図 10 の符号 501 に示すように、リードやライトのアクセス種別等を示すメモリコマンド 502 とそれに付随するデータ 503 とから構成される。データ 503 は一般的にライトデータであり、従ってリード命令のようにデータ 503 が付随しない場合もある。通信回路 109A は、このようなメモリアクセス命令 501 を受けると、図 10 の符号 511 に示すように、メモリコマンド 502 及びデータ 503 をそのままそっくり通信データ 512 として含み、それにネットワークコマンド 513、送付先 ID 514 及び送付元 ID 515 を付加した通信メッセージに変換して通信ケーブル 111 に出力する。送付先 ID 514 には、当該メモリアクセス命令を送るノードのノード番号及びノード内番号が設定され、送付元 ID 515 には自ノードのノード番号及びノード内番号が設定される。

### 【0074】

アクセス先のノードから返されてくるメモリアクセス命令に対する通信メッセージは、図 10 の符号 521 に示すようにネットワークコマンド 523、送付先 ID 524、送付元 ID 525 及び通信データ 522 から構成され、形式的には通信メッセージ 511 と同じであるが、通信データ 522 の部分には、先に発行したメモリアクセス命令の応答が含まれている。通信回路 109A では、通信メ

ッセージ 5 2 1 を受信すると、その通信データ 5 2 2 を抽出して図 1 0 の符号 5 3 1 に示すような形式の応答に変換してノード間通信バス 1 1 2 に出力する。応答 5 3 1 は、通信データ 5 2 1 に含まれていた応答コマンド 5 3 2 及びそれに付随するデータ 5 3 3 から構成される。

## 【 0 0 7 5 】

他方、他のノードから送られてくるメモリアクセス命令を含む通信メッセージは図 1 0 の符号 5 1 1 の通信メッセージと同じ形式であり、これを受信したとき、通信回路 1 0 9 A は、メモリアクセス命令 5 0 1 に変換してノード間通信バス 1 1 2 に出力する。また、このメモリアクセス命令 5 0 1 に対してノード間通信バス 1 1 2 から受信する応答は図 1 0 の符号 5 3 1 の応答と同じ形式であり、これを受信したとき、通信回路 1 0 9 A は、通信メッセージ 5 2 1 に変換して通信ケーブル 1 1 1 に出力する。

## 【 0 0 7 6 】

以上のような機能を持つ通信回路 1 0 9 A の構成例を図 1 1 に示す。この例の通信回路 1 0 9 A が図 3 の通信回路 1 0 9 と相違するところは、メモリアクセス用の変換部 1 6 6 を有することと、変換部が 1 6 1、1 6 2、1 6 6 の 3 つあるため、選択部 1 6 3 A 及び振分部 1 6 4 A の機能が拡張されており、また振分情報レジスタ 1 6 5 に変換部 1 6 6 への通信メッセージの振分情報が追加されている点にある。

## 【 0 0 7 7 】

メモリアクセス用の変換部 1 6 6 は、送付先 ID レジスタ 1 9 1、自 ID レジスタ 1 9 2、M→N 変換部 1 9 3 及び N→M 変換部 1 9 4 とで構成される。M→N 変換部 1 9 3 は、ノード間通信バス 1 1 2 から自ノードの CPU から出力された図 1 0 のメモリアクセス命令 5 0 1 を受信すると、図 1 0 の通信メッセージ 5 1 1 に変換して選択部 1 6 3 A を通じて通信ケーブル 1 1 1 に出力する。通信メッセージ 5 1 1 の送付先 ID 5 1 4 には、メモリアクセス命令 5 0 1 のメモリアドレスから求まるノード番号とメモリアクセス用変換部を指定するノード内番号とを設定し、送付元 ID 5 1 5 には、自 ID レジスタ 1 9 2 に事前に設定された自ノード番号及び変換部 1 6 6 を示すノード内番号を設定する。N→M 変換部 1

94は、他ノードから送られてきた図10の通信メッセージ521を振分部164Aから受信すると、図10のメモリアクセス命令531に変換してノード間通信バス112に出力する。

#### 【0078】

また、N→M変換部194は、他ノードから送られてきた図10の通信メッセージ511を振分部164Aから受信すると、図10のメモリアクセス命令501に変換してノード間通信バス112に出力する。このとき、通信メッセージ511の送付元ID515を当該メモリアクセス命令の応答を返す送付先IDとして送付先IDレジスタ191に格納しておく。ノード間通信バス112から当該メモリアクセス命令の応答をM→N変換部193が受信すると、図10の応答531から通信メッセージ521を生成するが、その送付先ID524に送付先IDレジスタ191に格納された送付先IDを設定する。

#### 【0079】

このように本例の計算機システムは、図8に示したように複数のCPU・メモリ搭載装置101-1A～101-3Aと複数の入出力制御装置201-1～201-3とを相互接続するネットワーク301を使って、CPU・メモリ搭載装置101-1A～101-3A間の通信を行うため、ネットワーク301の資源を有効に利用することが可能になる。

#### 【0080】

図12は本発明を適用した計算機システムの更に別の例を示すブロック図である。この例の計算機システムは、図1または図9に示したCPU・メモリ搭載装置101または101A及び図4に示した入出力制御装置201を、それぞれ601-1～601-3、201-1～201-3で示す各3台ずつ使用して、図6または図8と同様に3つの情報処理装置602～604を有するクラスタ型の計算機システムを実現すると共に、CPU・メモリ搭載装置601-1～601-3及び入出力制御装置201-1～201-3と同様なCPU・メモリ搭載装置601-4及び入出力制御装置201-4をそれぞれ1台ずつ、予備の装置として事前にネットワーク301に接続した構成を有する。なお、予備のCPU・メモリ搭載装置601-4及び入出力制御装置201-4には、診断制御装置3

1 4 が接続され、この診断制御装置 3 1 4 は他の診断制御装置 3 1 1 ~ 3 1 3 と同じく診断用ネットワーク 3 1 4 に接続されている。

【 0 0 8 1 】

この例の計算機システムでは、CPU・メモリ搭載装置 6 0 1 - 4 は通常システム運用時は使用されず、待機系として用意されており、他の現用系の CPU・メモリ搭載装置 6 0 1 - 1 ~ 6 0 1 - 3 に障害が発生して動作不能になった場合、動作不能となった CPU・メモリ搭載装置が使用していた入出力制御装置を引き取って運用を継続する。例えば、図 1 2 の CPU・メモリ搭載装置 6 0 1 - 1 が動作不能になると、診断制御装置 3 1 1 によって入出力制御装置 2 0 1 - 1 における通信回路 2 0 6 の相手 ID が CPU・メモリ搭載装置 6 0 1 - 4 の ID に再設定される等の処理が行われ、また診断制御装置 3 1 4 により CPU・メモリ搭載装置 6 0 1 - 4 の立ち上げが行われ、その際に CPU・メモリ搭載装置 6 0 1 - 4 の通信回路 1 0 9 または 1 0 9 A のポート 0 用の変換部 1 6 1 における相手 ID が入出力制御装置 2 0 1 - 1 の ID に設定される等の初期設定が行われ、CPU・メモリ搭載装置 6 0 1 - 1 で実施されていた業務処理が、入出力制御装置 2 0 1 - 1 を使って CPU・メモリ搭載装置 6 0 1 - 4 で再開される。

【 0 0 8 2 】

また、入出力制御装置 2 0 1 - 4 は通常システム運用時は使用されず、待機系として用意されており、他の入出力制御装置 2 0 1 - 1 ~ 2 0 1 - 3 に障害が発生して動作不能になった場合、動作不能となった入出力制御装置を使用していた CPU・メモリ搭載装置に割り当てて運用を継続する。例えば、図 1 2 の入出力制御装置 2 0 1 - 1 が動作不能になると、診断制御装置 3 1 1 によって CPU・メモリ搭載装置 6 0 1 - 1 の再立ち上げ時に通信回路 1 0 9 または 1 0 9 A のポート 0 用の変換部 1 6 1 における相手 ID が入出力制御装置 2 0 1 - 4 の ID に設定される等の初期設定が行われ、また診断制御装置 3 1 4 により入出力制御装置 2 0 1 - 1 における通信回路 2 0 6 の相手 ID が CPU・メモリ搭載装置 6 0 1 - 1 の ID に設定される等の処理が行われ、入出力制御装置 2 0 1 - 1 を使って実施されていた業務が入出力制御装置 2 0 1 - 4 を使って CPU・メモリ搭載装置 6 0 1 - 1 で再開される。

## 【0083】

以上本発明を幾つかの例を挙げて説明したが、本発明は以上の例に限定されず以下に述べるような各種の付加変更が可能である。

## 【0084】

CPU・メモリ搭載装置は、4台のCPU103を備えるようにしたが、4台以上でも、3台以下でも良く、少なくとも1台のCPUを備えていれば良い。

## 【0085】

CPU・メモリ搭載装置において、CPU103はCPUバス102を介して制御回路104に接続されていたが、制御回路104に個別の信号線によって接続されていても良いし、共通バスを介してCPU103、メモリ106、制御回路104及び通信回路109が相互に接続される構成であっても良い。

## 【0086】

CPU・メモリ搭載装置は、最大2台の入出力制御装置を接続する構成としたが、3台以上の入出力制御装置を接続できる構成にしても良く、また1台しか接続できない構成であっても良い。

## 【0087】

入出力制御装置は、1つの入出力制御回路202だけを備えていたが、2つ以上の入出力制御回路を備えていても良い。この場合、通信回路206は、図5に示した相手IDレジスタ221、自IDレジスタ222、N→I/O変換部223及びI/O→N変換部224を、入出力制御回路の数だけ備え、通信回路109と同様な選択部163、振分部164及び振分情報レジスタ165を備える構成となる。

## 【0088】

1台のCPU・メモリ搭載装置には、システムの通常運用時に1つの入出力制御装置だけを割り当てたが、システムの通常運用時に1台のCPU・メモリ搭載装置に複数の入出力制御装置を割り当てるようにしても良い。

## 【0089】

図7の最小構成の計算機システムは図1のCPU・メモリ搭載装置101を使ったが、図9のCPU・メモリ搭載装置101Aを使用することもできる。



## 【0090】

図2で、通信回路109で行われる変換処理の概要で、I/Oコマンド122とデータ123をそのままそっくり通信データ132に含むように変換したが、I/Oコマンドと一対一のネットワークコマンドを定義することで、I/Oコマンドとネットワークコマンドを統合させたり、データ自体は圧縮してネットワークデータとして送るといような手法を取ることも出来る。また、送付先IDや送付元IDとして、ノード番号と入出力ポートを一意に識別するために上位数ビット、下位数ビットを設定するようにしているが、入出力ポートをシステムで一意に決定する番号であればよく、入出力ポート単位に連続した値をとっても良いし、任意の一意的な値と定義しても良い。

## 【0091】

## 【発明の効果】

以上説明したように本発明によれば、以下のような効果が得られる。

## 【0092】

計算機システムにおける障害時の可用性をより一層高めることができる。その理由は、従来の計算機システムでは、CPU及びメモリ部分に障害が発生して使用不能になると、それに直結されていた入出力制御装置も、たとえ正常であっても使用できなくなるが、本発明では、CPU・メモリ搭載装置と入出力制御装置とが分離しており、入出力制御装置を別のCPU・メモリ搭載装置で使用することができるからである。また、従来の計算機システムでは、入出力制御装置に障害が発生して使用不能になると、それに直結されていたCPU及びメモリがそれ以外に入出力制御装置を有しない場合、たとえ正常であっても使用できなくなるが、本発明では、予備の入出力制御装置を当該CPU・メモリ搭載装置に割り当てることで、当該CPU・メモリ搭載装置の使用が可能になるからである。

## 【0093】

計算機システムのオペレーティングシステムを改造する必要がない。その理由は、通信手段がCPUから発行された入出力命令の所定の宛先への送信とその応答の受信を担っているため、CPUは入出力命令の発行やその応答の受け取りに関して従来と何ら変わるところがなく、入出力制御装置があたかも直結されてい

るかのように見えるからである。

【 0 0 9 4 】

ネットワーク接続時に問題となり易い予期せぬ相手からの通信による誤動作を防止することができる。その理由は、入出力制御装置の通信手段は、ネットワーク経由で受信した入出力命令の送付元が事前に設定されたCPU・メモリ搭載装置である場合に限って当該入出力命令を有効なものとして受信する手段を備えているからである。また、CPU・メモリ搭載装置の通信手段は、ネットワーク経由で受信した応答の送付元が事前に設定された入出力制御装置である場合に限って当該応答を有効なものとして受信する手段を備えているからである。

【図面の簡単な説明】

【図 1】

本発明を適用した計算機システムで使用されるCPU・メモリ搭載装置の一例を示すブロック図である。

【図 2】

CPU・メモリ搭載装置及び入出力制御装置の通信回路で行われる変換処理の概要を示す図である。

【図 3】

CPU・メモリ搭載装置の通信回路の構成例を示すブロック図である。

【図 4】

本発明を適用した計算機システムで使用される入出力制御装置の一例を示すブロック図である。

【図 5】

入出力制御装置の通信回路の構成例を示すブロック図である。

【図 6】

本発明を適用した計算機システムの一例を示すブロック図である。

【図 7】

本発明を適用した計算機システムの別のブロック図である。

【図 8】

本発明を適用した計算機システムの更に別のブロック図である。

【図 9】

本発明を適用した計算機システムで使用される CPU・メモリ搭載装置の別の例を示すブロック図である。

【図 1 0】

CPU・メモリ搭載装置の通信回路で行われるメモリアクセス関連の変換処理の概要を示す図である。

【図 1 1】

CPU・メモリ搭載装置の通信回路の別の構成例を示すブロック図である。

【図 1 2】

本発明を適用した計算機システムの更に別の例を示すブロック図である。

【図 1 3】

従来の計算機システムの一例を示すブロック図である。

【図 1 4】

従来の計算機システムの別の例を示すブロック図である。

【符号の説明】

- 1 0 1 … CPU・メモリ搭載装置
- 1 0 2 … CPUバス
- 1 0 3 … CPU
- 1 0 4 … 制御回路
- 1 0 5 … メモリ信号線
- 1 0 6 … メモリ
- 1 0 7、1 0 8 … 入出力信号線
- 1 0 9 … 通信回路
- 1 1 0 … 接続具
- 1 1 0 … 通信ケーブル
- 2 0 1 … 入出力制御装置
- 2 0 2 … 入出力制御回路
- 2 0 3 … 標準入出力バス
- 2 0 4 … 入出力カード

2 0 5 … 入出力信号線

2 0 6 … 通信回路

2 0 7 … 接続具

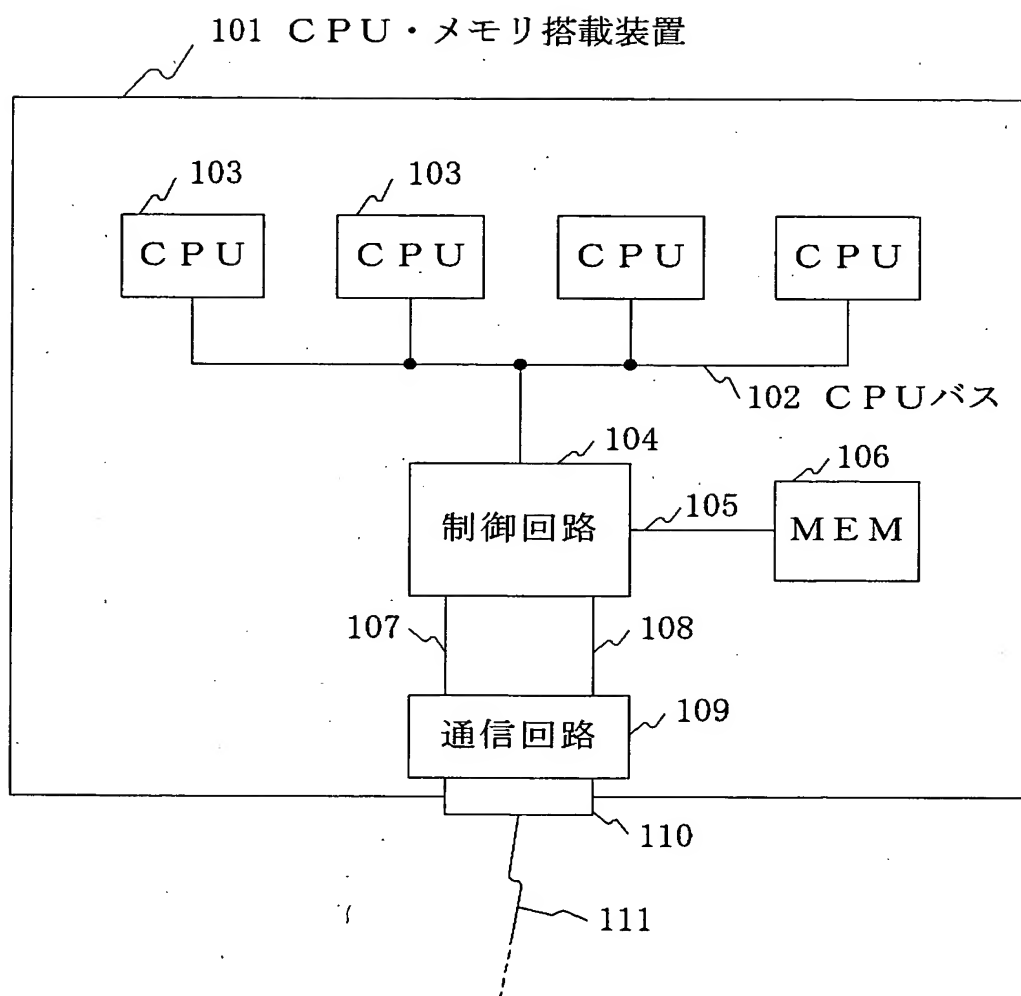
2 1 1 … 通信ケーブル

2 1 2 … 周辺機器

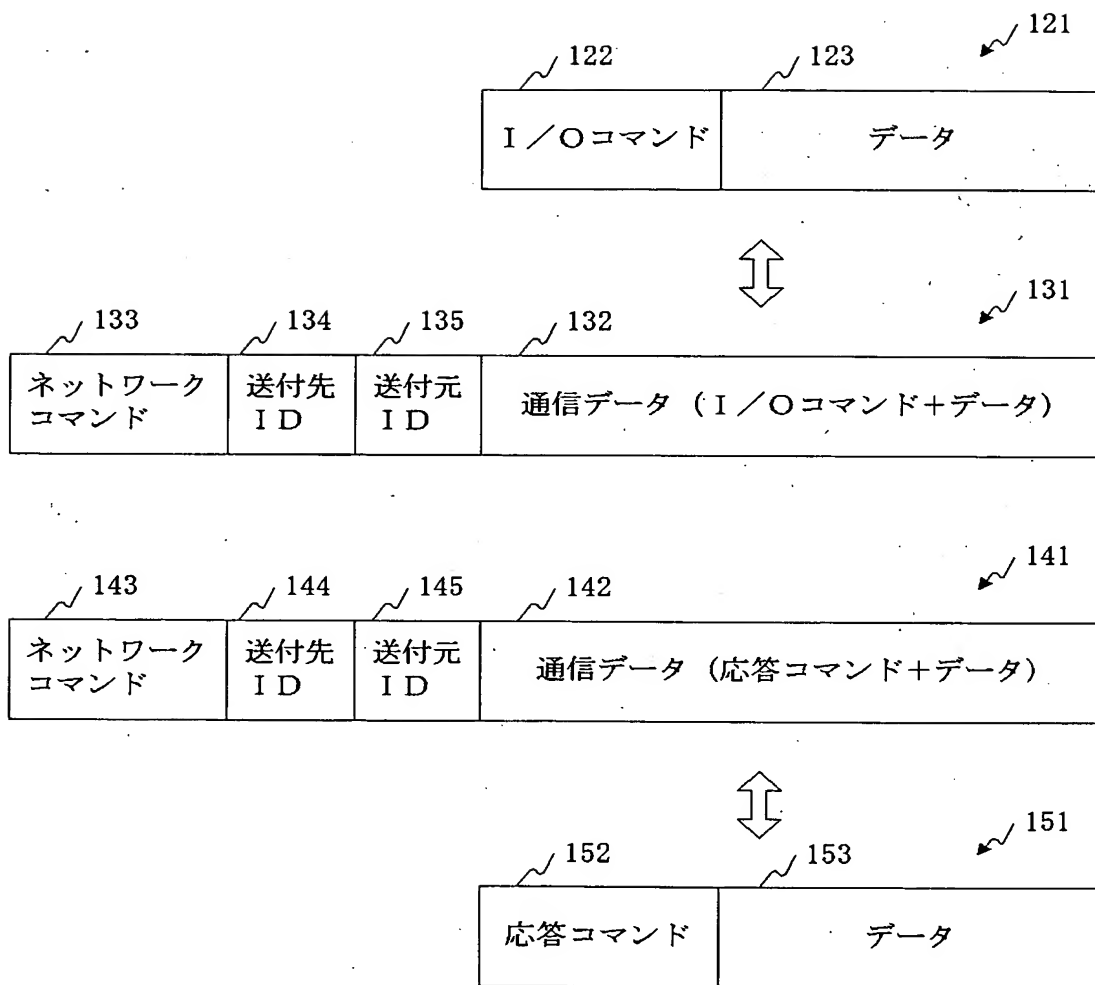
3 0 1 … ネットワーク

【書類名】 図面

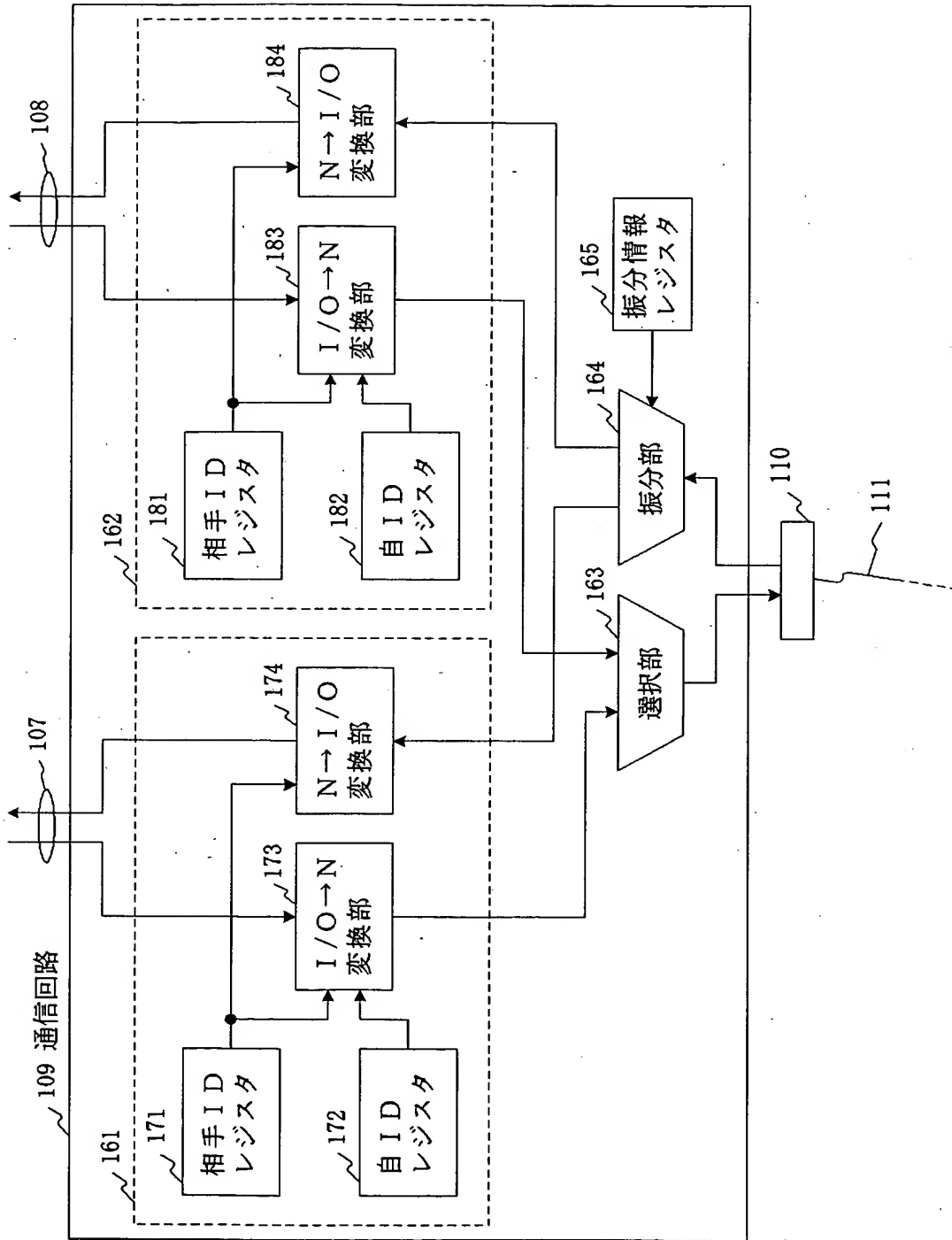
【図1】



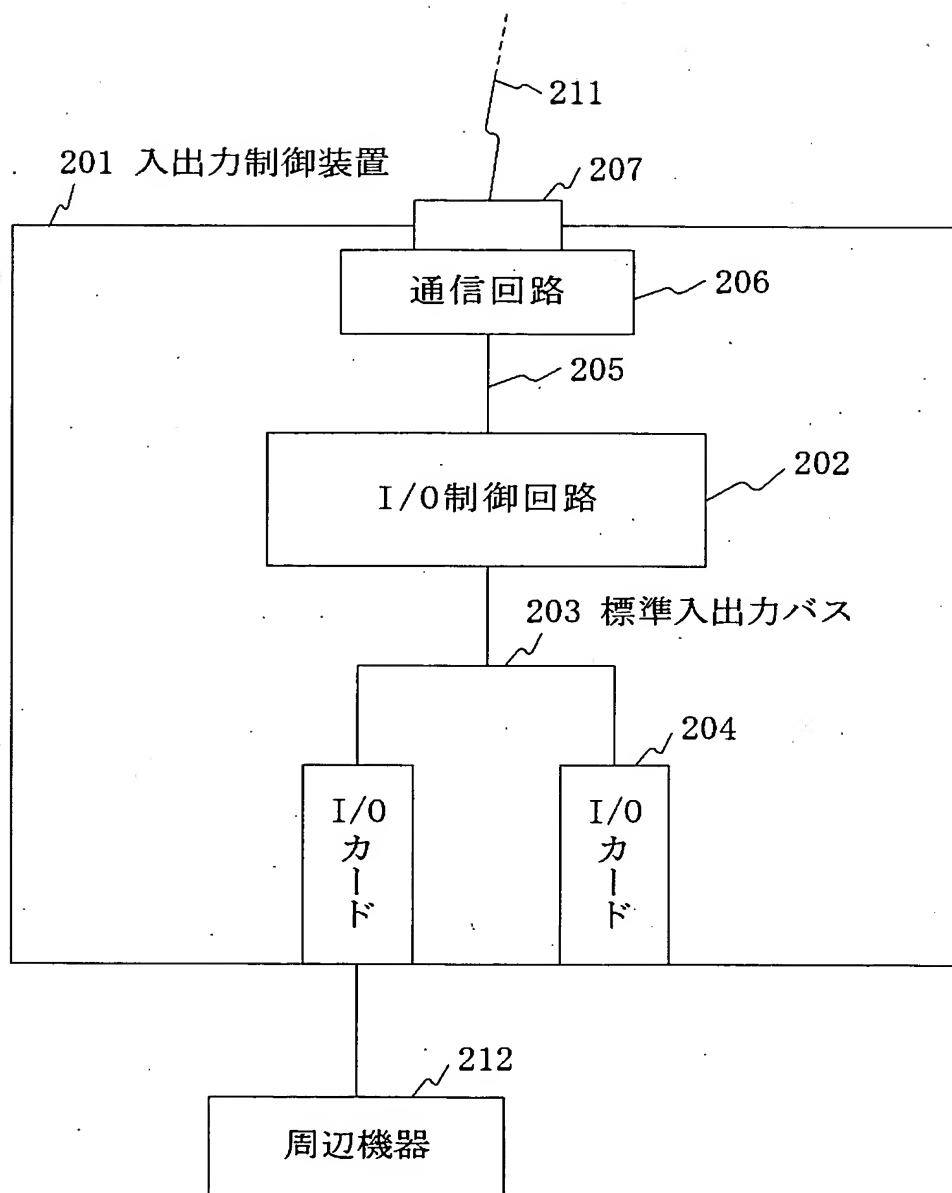
【図 2】



【図 3】

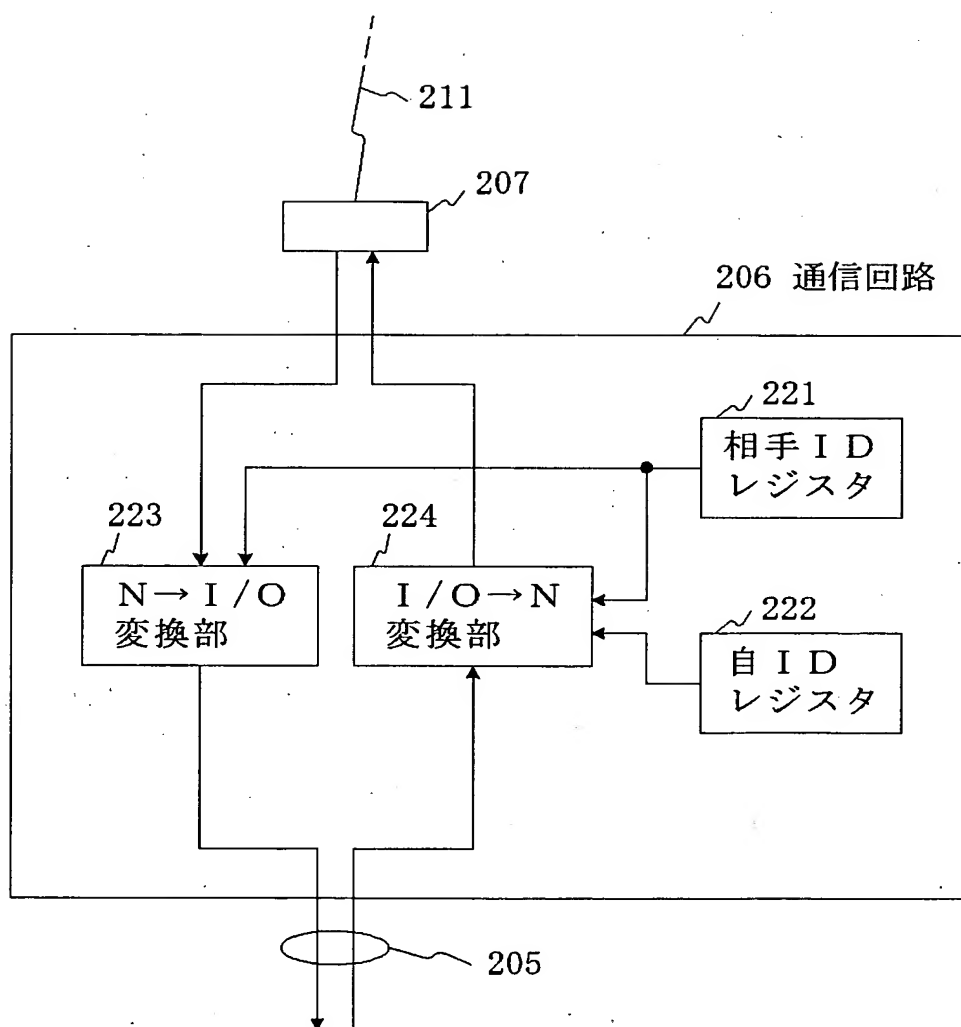


【図 4】

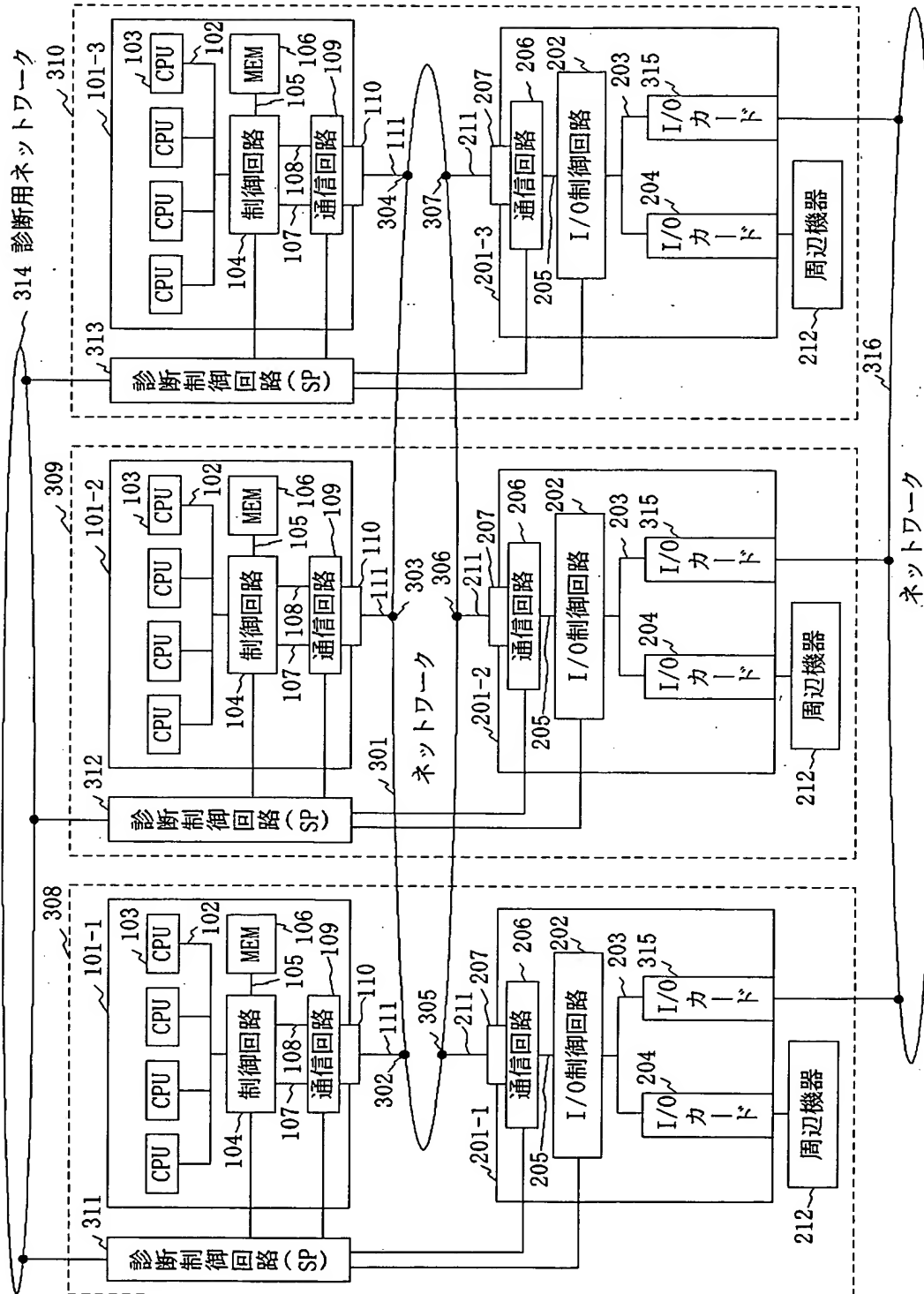




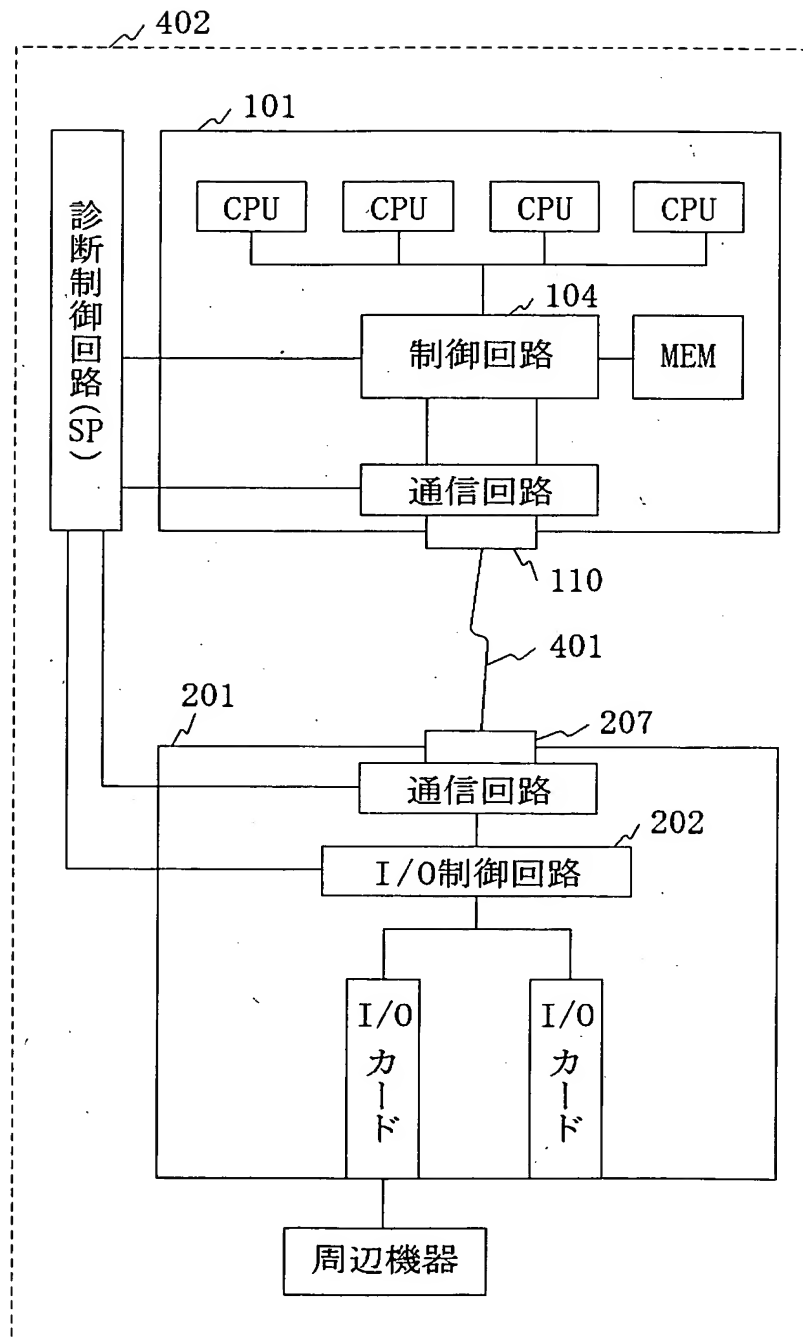
【図5】



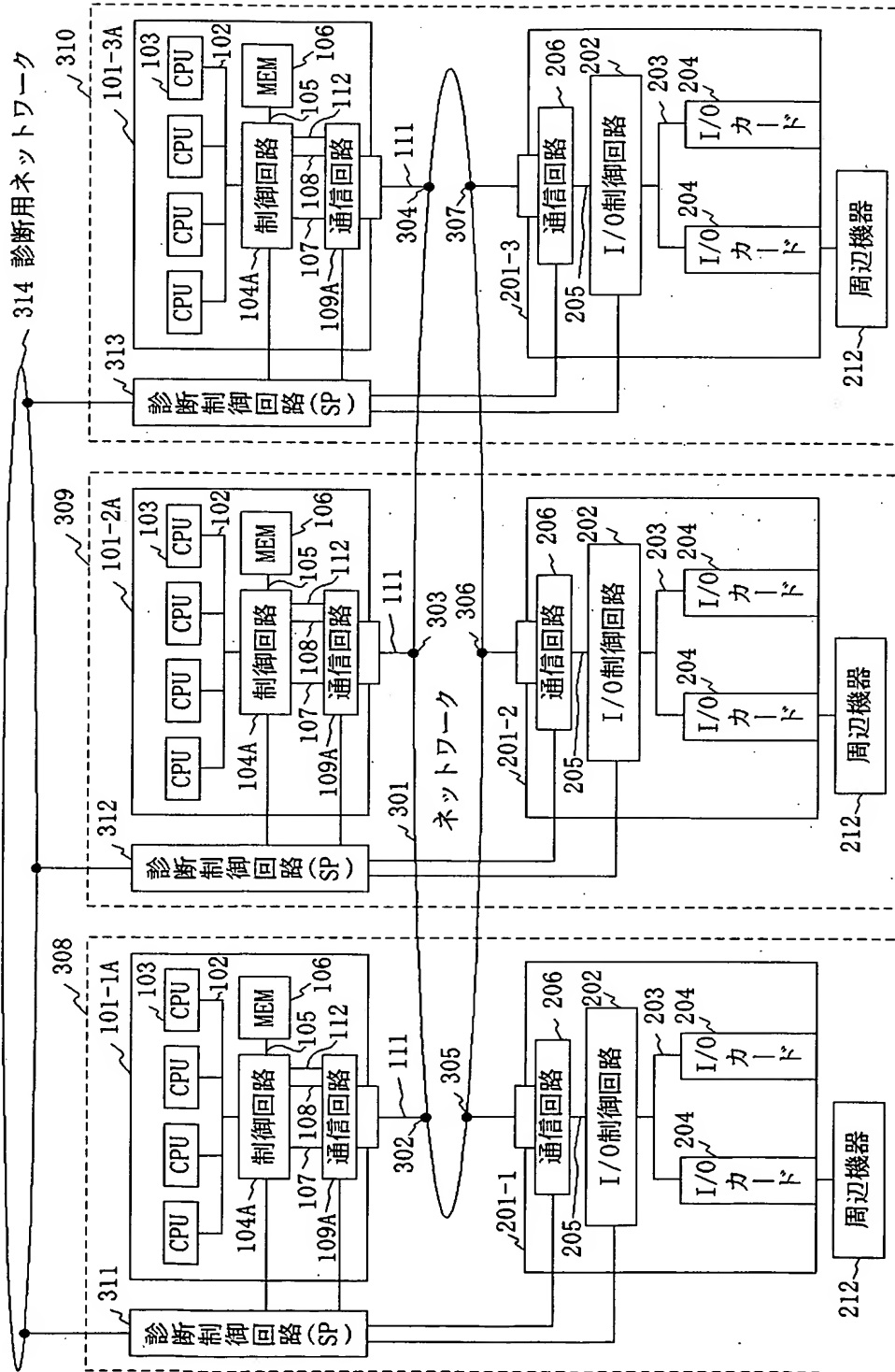
【図 6】



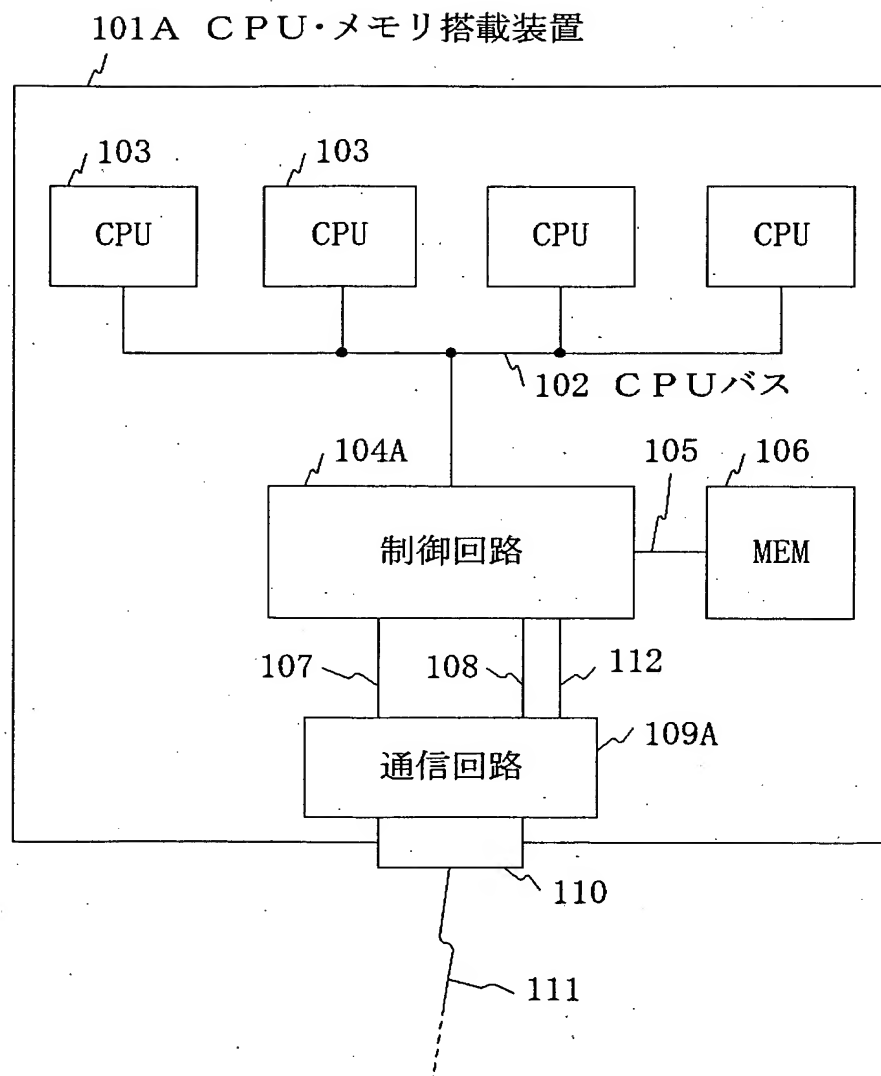
【図 7】



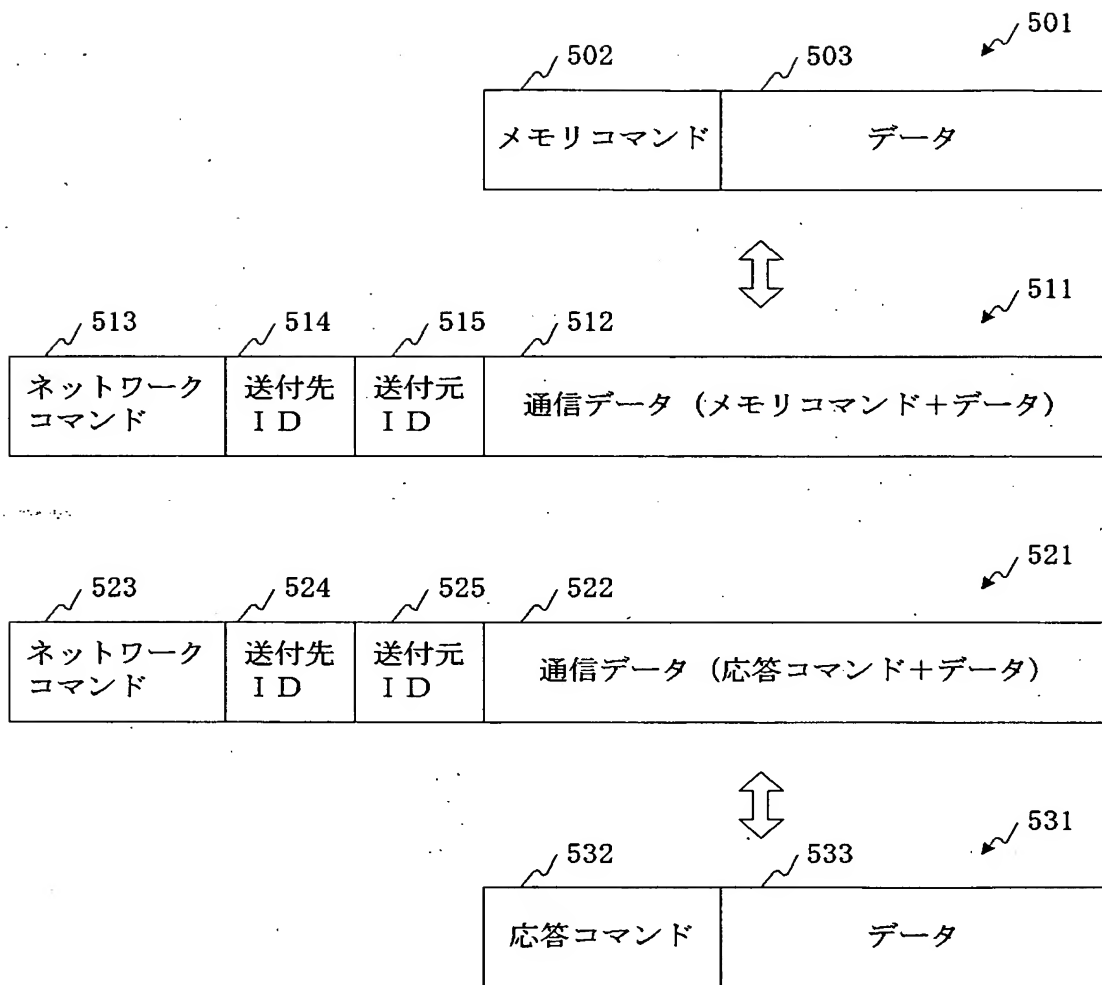
【図 8】



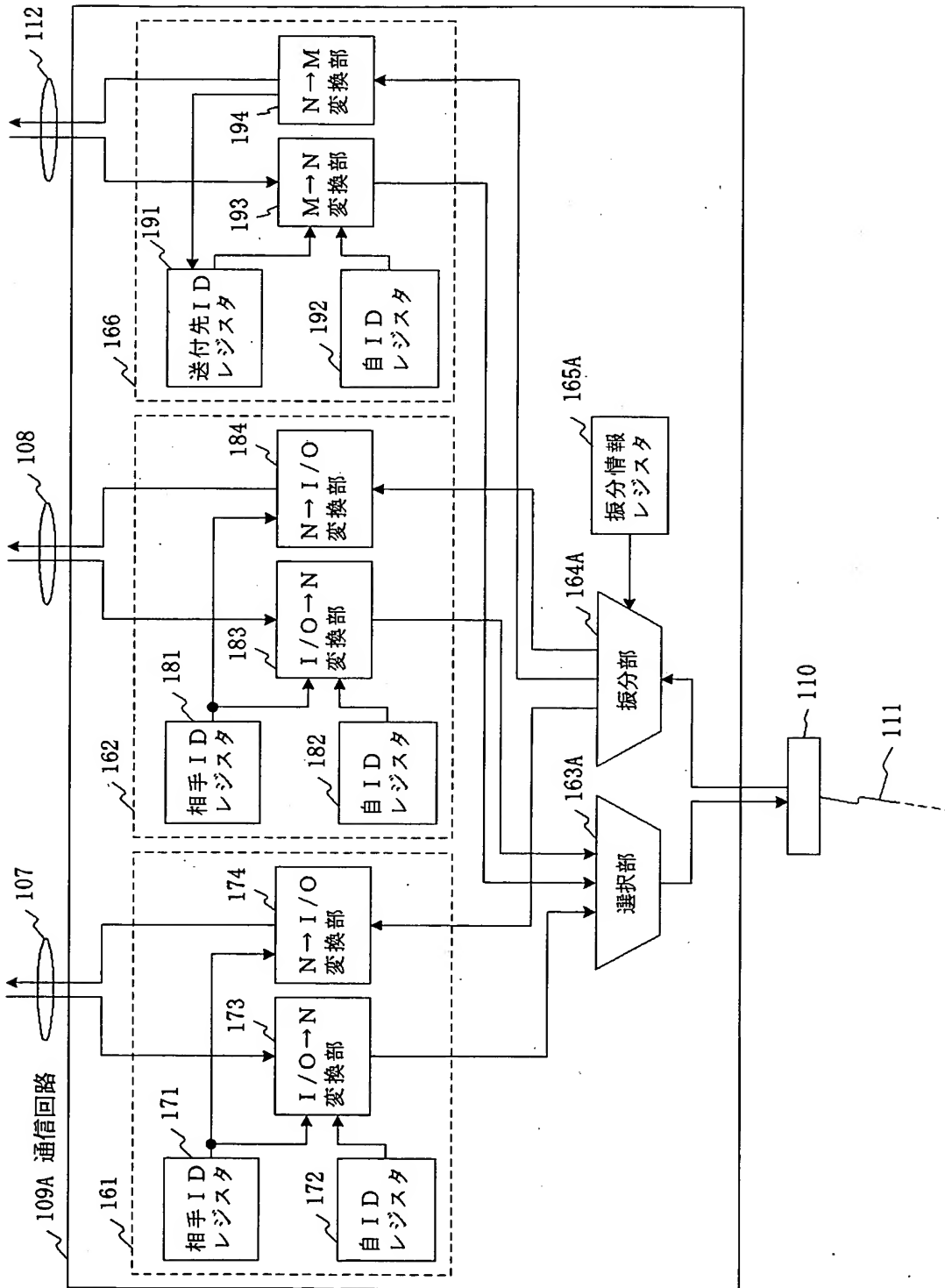
【図9】



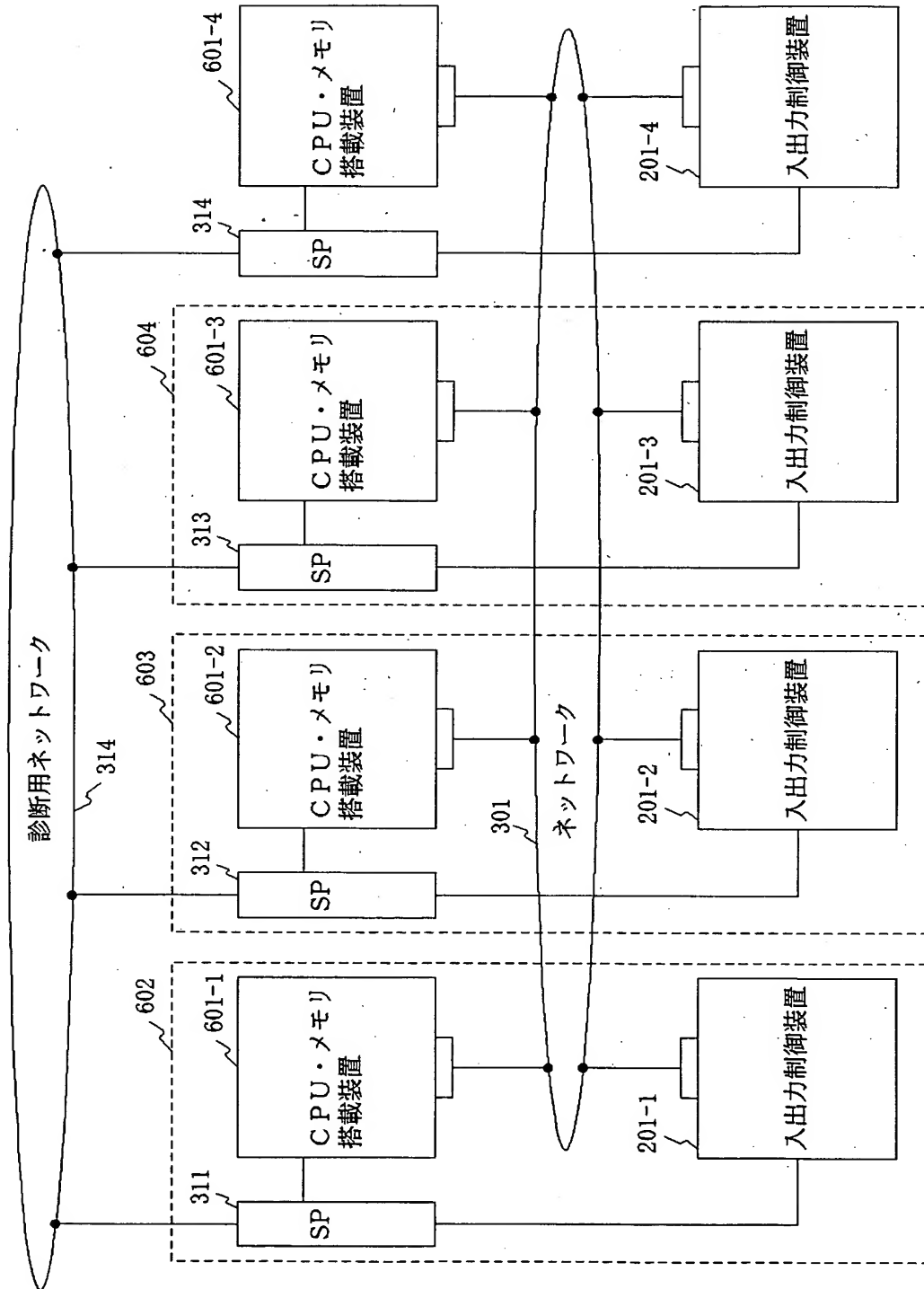
【図10】



【図 11】

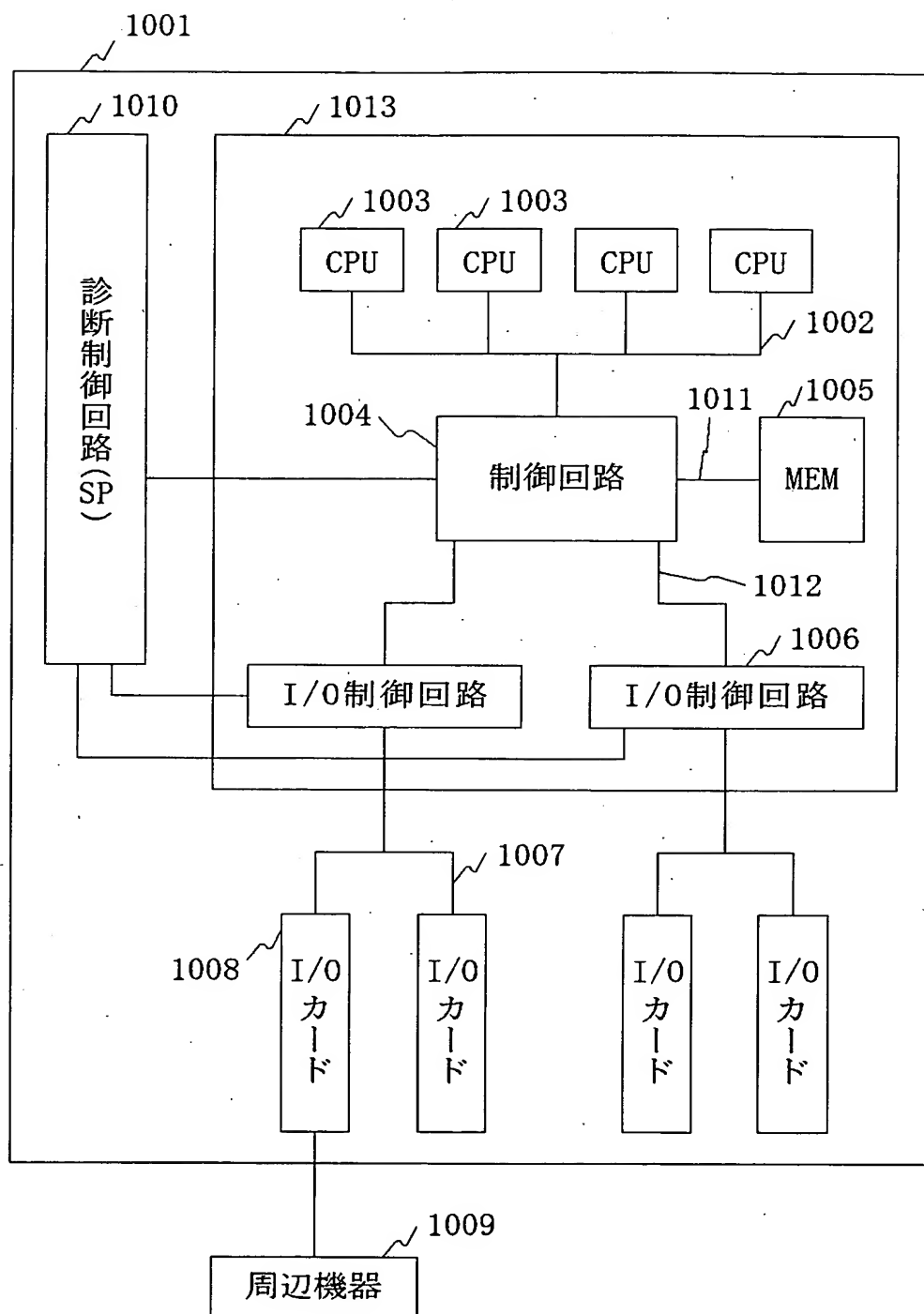


【図 12】

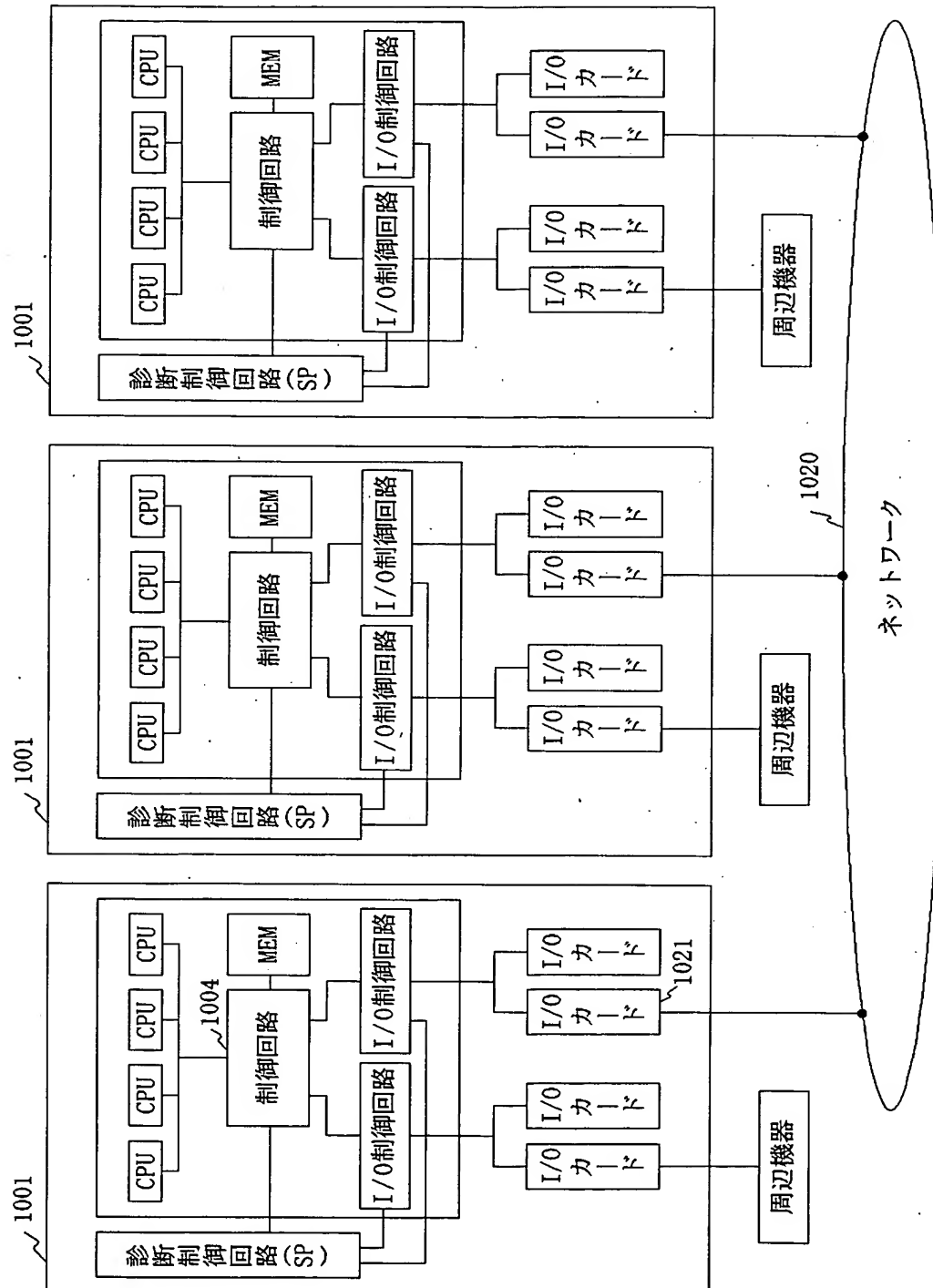




【図13】



【図14】



【書類名】 要約書

【要約】

【課題】 計算機システムの可用性を向上させる。

【解決手段】 CPU103 及びメモリ106 を有するCPU・メモリ搭載装置101-1 ～101-3 と、入出力制御装置201-1 ～201-3 とをネットワーク301 により相互に接続する。CPU・メモリ搭載装置101-1 の通信回路109 は、自装置のCPU103 から発行された入出力命令を自装置に事前に割り当てられた入出力制御装置201-1 に対してネットワーク301 経由で送信し、その応答をネットワーク301 経由で受信する。入出力制御装置201-1 の通信回路206 は、自装置に事前に割り当てられたCPU・メモリ搭載装置101-1 からの入出力命令をネットワーク301 経由で受信し、その応答をネットワーク301 経由で送信する。CPU・メモリ搭載装置101-1 が障害により使用不能になると、入出力制御装置201-1 は他の正常なCPU・メモリ搭載装置101-2 に割り当てられる。

【選択図】 図6

出 願 人 履 歷 情 報

識別番号 [000004237]

1. 変更年月日 1990年 8月29日

[変更理由] 新規登録

住 所 東京都港区芝五丁目7番1号

氏 名 日本電気株式会社